CHL 5225H
Advanced Statistical Methods for Clinical Trials: Survival Analysis

Prof. Kevin E. Thorpe

Dalla Lana School of Public Health
University of Toronto

Outline

Defining Survival Data

Introduction

Interest centres on patients who are at risk for some event and the time to the occurrence of the event.

Examples: Time to death
Time to stroke
Time to disease recurrence

Definition of Failure Time

Three requirements to precisely determine failure time:

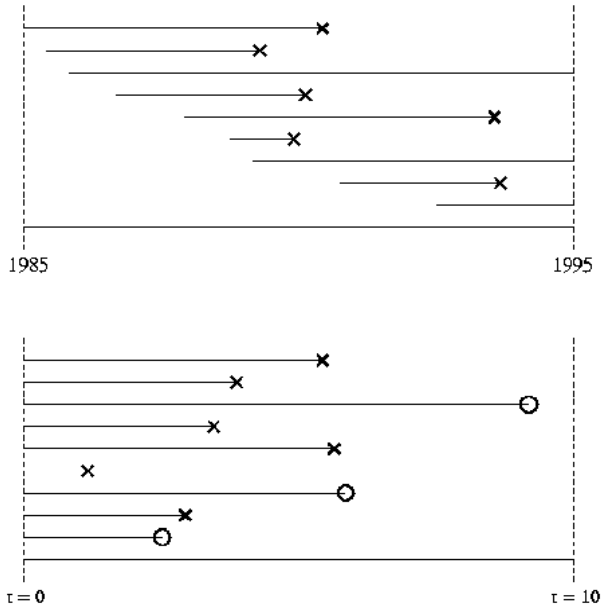
- unambiguous time origin
- scale for measuring the passage of time
- meaning of the failure event must be entirely clear

Censoring

Patients who are not observed to "failure" are called *censored*
Some Causes:

- study concludes before all patients destined to fail have actually failed
- patient withdraws from study before failure observed

Typical Survival Data Arising From a Clinical Trial



Mathematical Definitions

Introduction

- Let the "failure time" be represented by a non-negative random variable T.
- Usually, we consider continuous distributions.
- Discrete and mixed discrete-continuous distributions can also handled.

The Survivor Function

- The *survivor function* $\mathcal{F}(t)$ is defined to be

$$\mathcal{F}(t) = \Pr(T \geq t)$$

- It is sometimes useful to consider the cumulative risk of an event. The *cumulative risk function* is defined by

$$R(t) = 1 - \mathcal{F}(t)$$

- Note that the definition of $\mathcal{F}(t)$ leads to left continuous CDF rather than the usual right continuity.

Probability Density Function

- Continuous T.

$$f(t) = -\mathcal{F}'(t) = \lim_{\Delta \to 0+} \frac{\Pr(t \leq T < t + \Delta)}{\Delta}$$

which gives

$$\mathcal{F}(t) = \int_t^\infty f(u)du$$

- Discrete T.
Usually handled by assigning to the density a component $f_j \delta(t - a_j)$ for an atom f_j at a_j where $\delta(\cdot)$ is the Dirac delta function.

Hazard Function

Continuous Distributions

- The *hazard function* (also called age-specific failure rate or force of mortality is defined by

$$h(t) = \lim_{\Delta \to 0+} \frac{\Pr(t \leq T < t + \Delta | t \leq T)}{\Delta}$$

- By the definition of conditional probability

$$h(t) = f(t)/\mathcal{F}(t)$$

- Sometimes the hazard is denoted by $\lambda(t)$.

Hazard Function

Continuous Distributions (continued)

- By the definitions of $f(t)$ and that $\mathcal{F}(0) = 1$ we have

$$\begin{aligned} h(t) &= -\mathcal{F}'(t)/\mathcal{F}(t) \\ &= -d \log \mathcal{F}(t)/dt \end{aligned}$$

and

$$\begin{aligned} \mathcal{F}(t) &= \exp\left(-\int_0^t h(u)du\right) \\ &= \exp[-H(t)] \end{aligned}$$

where $H(\cdot)$ is called the integrated hazard.

- Finally,

$$f(t) = h(t)exp[-H(t)]$$

Hazard Function

Discrete Distributions

- If there is an atom f_j of probability at time a_j then

$$\begin{aligned} h_j &= f_j/\mathcal{F}(a_j) \\ &= f_j/(f_j + f_{j+1} + \cdots) \end{aligned}$$

- With this definition it can be shown that

$$\mathcal{F}(t) = \prod_{a_j < t} (1 - h_j)$$

- By defining the integrated hazard as

$$H(t) = \sum_{a_j < t} \log(1 - h_j)$$

we still have

$$\mathcal{F}(t) = \exp[-H(t)]$$

The Likelihood Function

We assume each patient has a failure time t_i and a censoring time c_i associated with them. In practice, we observe $x_i = \min(t_i, c_i)$.

$$\begin{aligned} \mathcal{L} &= \prod_u f(t_i;\phi)\prod_c \mathcal{F}(c_i;\phi) \\ \ell &= \sum_u \log f(t_i;\phi) + \sum_c \log \mathcal{F}(c_i;\phi) \\ \ell &= \sum_u \log h(x_i;\phi) + \sum \log \mathcal{F}(x_i;\phi) \end{aligned}$$

Survival Models

- Suppose in addition to survival some covariates z are measured with the intent of modeling their association with survival.
- For some "baseline" hazard function $h_0(t)$ (or survivor function $\mathcal{F}_0(t)$ we have the following commonly used models.
  - The *Accelerated Life* model is defined by

    $$\mathcal{F}(t;z) = \mathcal{F}_0[t\psi(z)] \text{ or } h(t;z) = h_0[t\psi(z)]\psi(z)$$

    The choice of $\mathcal{F}_0$ is dictated by a choice of parametric model.
  - The *Proportional Hazard* model is defined by

    $$\mathcal{F}(t;z) = [\mathcal{F}_0(t)]^{\psi(z)} \text{ or } h(t;z) = \psi(z)h_0(t)$$

    The baseline hazard does not need to be explicitly known for this model.

The Likelihood Function

- We begin by assuming a discrete distribution atoms of probability at g unique failure times $t_1, t_2, \ldots, t_g$.
- It can be shown that the log-likelihood is

  $$\ell = \sum_j [d_j \log h_j + (r_j - d_j)\log(1 - h_j)]$$

  where r_j is the number of patients in view at time t_j and d_j is the number who fail at t_j.
- By convention, r_j includes individuals who are censored at t_j.
- This is identical to the log likelihood for g independent binomials with r_j trials, d_j failures and probability of failure h_j.
- This fact is used in the derivation of the results on the next slide.

Product-Limit or Kaplan-Meier Estimate

- Suppose g unique failure times are observed. Put them in order from smallest to largest.

  $$t_1, t_2, \ldots, t_g$$

- The survivor function is estimated at time t_j by

  $$\hat{\mathcal{F}}(t_j) = \prod_{i=1}^{j}\left(1 - \frac{d_i}{r_i}\right) = \prod_{i=1}^{j}\left(\frac{r_i - d_i}{r_i}\right)$$

- Note the similarity between the first expression above and the expression for a discrete survivor function, $\mathcal{F}(t) = \prod_{a_j<t}(1 - h_j)$ given previously.
- The variance of $\hat{\mathcal{F}}(t)$ at time t_j is (Greenwood's formula)

  $$V(\hat{\mathcal{F}}(t_j)) = [\hat{\mathcal{F}}(t_j)]^2 \sum_{i=1}^{j} \frac{d_i}{r_i(r_i - d_i)}$$

Example

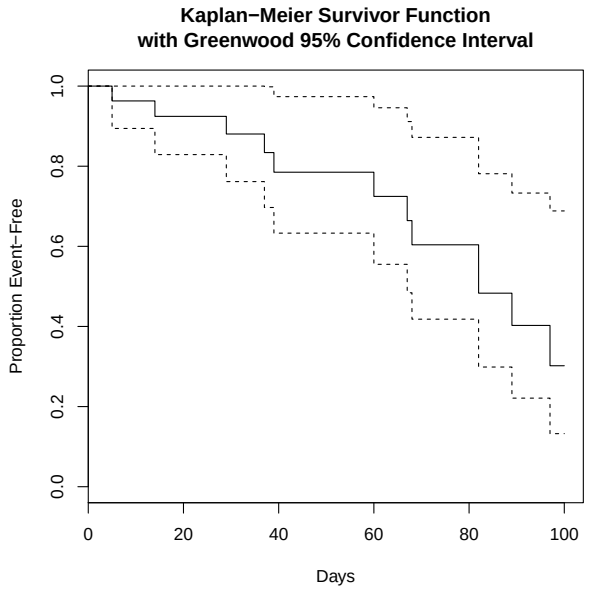
Suppose we had the following data (+ indicates censored observations).

```
 [1]   2+   2+   5    9+  14   16+  16+  17+  29   30+  37
[12]  37+  39   44+  44+  58+  60   67   68   82   82   86+
[23]  86+  89   93+  97   100+ 100+ 100+
```

Example

t_j	r_j	d_j	$P(t_j) = 1 - d_j/r_j$	$\hat{\mathcal{F}}(t_j)$	SE
0	29			1.0	
5	27	1	$1 - 1/27 = 0.9630$	$1.0 \times 0.9630 = 0.963$	0.0363
14	25	1	$1 - 1/25 = 0.9600$	$0.963 \times 0.9600 = 0.924$	0.0514
29	21	1	$1 - 1/21 = 0.9524$	$0.924 \times 0.9524 = 0.880$	0.0651
37	19	1	$1 - 1/19 = 0.9474$	$0.880 \times 0.9474 = 0.834$	0.0764
39	17	1	$1 - 1/17 = 0.9412$	$0.834 \times 0.9412 = 0.785$	0.0863
60	13	1	$1 - 1/13 = 0.9231$	$0.785 \times 0.9231 = 0.725$	0.0985
67	12	1	$1 - 1/12 = 0.9167$	$0.725 \times 0.9167 = 0.664$	0.1072
68	11	1	$1 - 1/11 = 0.9091$	$0.664 \times 0.9091 = 0.604$	0.1132
82	10	2	$1 - 2/10 = 0.8000$	$0.604 \times 0.8000 = 0.483$	0.1185
89	6	1	$1 - 1/6 = 0.8333$	$0.483 \times 0.8333 = 0.403$	0.1231
97	4	1	$1 - 1/4 = 0.7500$	$0.403 \times 0.7500 = 0.302$	0.1270

Example



Example

R Code

```
> library(survival)
> eg1a.km <- survfit(Surv(days, status), data = eg1a)
> summary(eg1a.km)
> plot(eg1a.km, mark.time = FALSE, xlab = "Days",
+      ylab = "Proportion Event-Free",
+      main = "Kaplan-Meier Survivor Function\n
with Greenwood 95% Confidence Interval")
```

Actuarial Estimator

- Suppose instead of exact failure/censoring times we know only the numbers of failures and censored observations between time intervals.
- For each interval $(t_{j-1}, t_j]$ we define r_{j-1} to be the number at risk at t_{j-1}, d_j to be the number who fail during the interval and m_j to be the number censored during the interval.
- Define $r'_j = r_{j-1} - \frac{1}{2}m_j$.
- Then, the actuarial estimator of survival is

$$\tilde{\mathcal{F}}(t_j) = \prod_{k \leq j} \left(1 - \frac{d_k}{r'_k}\right)$$

Comparing Survival

- There are a variety of methods that can be used to compare survival in various ways.
  - Point in time comparison (two groups)
  - Mantel-Haenszel or Log Rank Test (two or more survival curves)
  - Model-based tests (two or more groups)

Point in Time Comparison

- The object is to compare survival between two treatments at a pre-specified time point t_0.
- From the Kaplan-Meier procedure and Greenwood's formula we have,

$$\begin{aligned} \text{Treatment} &: \hat{\mathcal{F}}_T(t_0) \text{ and } V[\hat{\mathcal{F}}_T(t_0)] \\ \text{Control} &: \hat{\mathcal{F}}_C(t_0) \text{ and } V[\hat{\mathcal{F}}_C(t_0)] \end{aligned}$$

- Then, to conduct an approximate test of the null hypothesis $H_0 : \mathcal{F}_T(t_0) = \mathcal{F}_C(t_0)$ we calculate the Z statistic,

$$Z = \frac{\hat{\mathcal{F}}_T(t_0) - \hat{\mathcal{F}}_C(t_0)}{\sqrt{V[\hat{\mathcal{F}}_T(t_0)] + V[\hat{\mathcal{F}}_C(t_0)]}}$$

Under H_0, $Z \sim N(0, 1)$.

Mantel-Haenszel Test

- Consider the two survival curves as a series of 2x2 contingency tables constructed at each time point where one or more events occur. At a time point t_j we have a table of the form:

	Treatment	Control	Total
Dead	d_{T_j}	d_{C_j}	d_j
Alive	$n_{T_j} - d_{T_j}$	$n_{C_j} - d_{C_j}$	$N_j - d_j$
Total	n_{T_j}	n_{C_j}	N_j

Then, from the hypergeometric distribution,

$$\begin{aligned} E(d_{T_j}) &= \frac{n_{T_j} d_j}{N_j} \\ V(d_{T_j}) &= \frac{n_{T_j} n_{C_j} (N_j - d_j) d_j}{N_j^2 (N_j - 1)} \end{aligned}$$

Mantel-Haenszel Test — Continued

- Define the following quantities.

$$O_T = \sum d_{T_j} \qquad E_T = \sum E(d_{T_j})$$
$$O_C = \sum d_j - O_T \qquad E_C = \sum d_j - E_T$$
$$V = \sum V(d_{T_j})$$

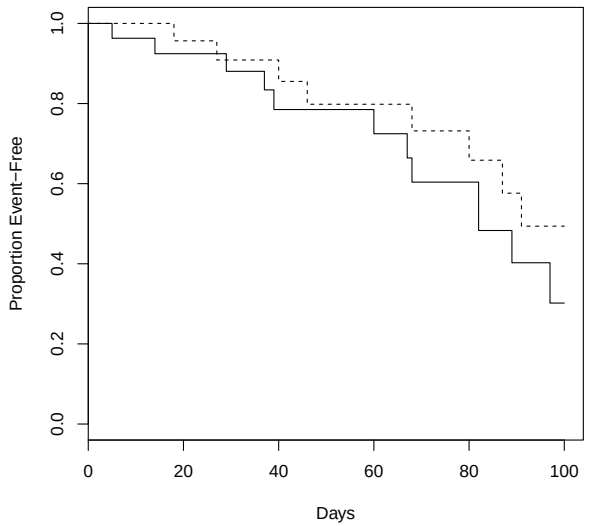
- Now, the null hypothesis $H_0 : \mathcal{F}_T(t) = \mathcal{F}_C(t)$ is usually tested with a $\chi^2_{(1)}$ test, but may also be tested with a Z test.

$$\chi^2 = \frac{(O_T - E_T)^2}{V}$$

and

$$Z = \frac{O_T - E_T}{\sqrt{V}}$$

- This is often called the *Log Rank Test*.



Comparing Survival in R

- In R the `survdiff()` function is used to compare survival curves.
- This performs tests in a family described by Harrington and Fleming (*Biometrika*: 1982) where the weights on each event time are $\mathcal{F}(t)^\rho$.
- The `survdiff()` function accepts an argument `rho`. If `rho = 0` (the default) you get the Mantel-Haenszel test and if `rho = 1` you get the Peto & Peto modification of the Gehan-Wilcoxon test.

Example

```
> survdiff(Surv(days, status) ~ rx, data = eg1b)

Call:
survdiff(formula = Surv(days, status) ~ rx, data = eg1b)

               N Observed Expected (O-E)^2/E (O-E)^2/V
rx=Control    29       12      9.9     0.447     0.898
rx=Treatment  25        8     10.1     0.438     0.898

 Chisq= 0.9  on 1 degrees of freedom, p= 0.343
```

Measuring the Treatment Effect

- Among other things, clinical trials are intended to provide an estimate of the treatment effect.
- For a "survival" outcome, two types of estimates are commonly used.
  - Point in time Risk Ratio or Risk Reduction.
  - All time or whole curve effect measures by means of Hazard Ratios.

Point in Time Treatment Effect

- Suppose a trial is comparing a treatment T to a control C.
- The usual expectation would be that the risk of the event would be lower in the treatment group than the control group.
- There may be situations where there is some clinically meaningful time point t_0 say, that is of interest.
- Consider three point in time estimates of "risk reduction."

$$\begin{aligned} \text{Risk Difference} &= R_C(t_0) - R_T(t_0) \\ \text{Risk Ratio} &= R_T(t_0)/R_C(t_0) \\ \text{Relative Risk Reduction} &= \frac{R_C(t_0) - R_T(t_0)}{R_C(t_0)} \\ &= 1 - \frac{R_T(t_0)}{R_C(t_0)} \end{aligned}$$

Estimating Risk Difference

- Estimate T & C survival curves using Kaplan-Meier procedure and compute.

$$\begin{aligned} \widehat{\text{RD}} &= \hat{R}_C(t_0) - \hat{R}_T(t_0) \\ &= (1 - \hat{\mathcal{F}}_C(t_0)) - (1 - \hat{\mathcal{F}}_T(t_0)) \\ &= \hat{\mathcal{F}}_T(t_0) - \hat{\mathcal{F}}_C(t_0) \end{aligned}$$

- Calculate $V(\hat{\mathcal{F}}_T(t_0))$ and $V(\hat{\mathcal{F}}_C(t_0))$ using Greenwood's formula (ie. SE^2).
- Then, 95% CI on the RD is (approximately):

$$\widehat{\text{RD}} \pm 1.96\sqrt{V(\hat{\mathcal{F}}_T(t)) + V(\hat{\mathcal{F}}_C(t))}$$

Estimating Risk Ratio

- Estimate T & C survival curves using Kaplan-Meier procedure and compute.

$$\begin{aligned} \widehat{\text{RR}} &= \hat{R}_T(t_0)/\hat{R}_C(t_0) \\ &= \frac{1 - \hat{\mathcal{F}}_T(t_0)}{1 - \hat{\mathcal{F}}_C(t_0)} \end{aligned}$$

- Calculate $V(\log \hat{R}_T(t_0))$ and $V(\log \hat{R}_C(t_0))$ using the following formula.

$$V(\log \hat{R}(t_0)) = \frac{V(\hat{\mathcal{F}}(t_0))}{(\hat{R}(t_0))^2}$$

- Then, 95% CI on RR is (approximately):

$$\widehat{\text{RR}} \times e^{\pm 1.96\sqrt{V(\log \hat{R}_T(t_0)) + V(\log \hat{R}_C(t_0))}}$$

Estimating Relative Risk Reduction

- Estimate T & C survival curves using Kaplan-Meier procedure and compute.

$$\begin{aligned}\widehat{\mathrm{RRR}} &= 1 - \frac{\hat{R}_T(t_0)}{\hat{R}_C(t_0)} \\ &= 1 - \widehat{\mathrm{RR}}\end{aligned}$$

- Determine the CI for $\widehat{\mathrm{RR}}$. Say this is $(\mathrm{RR}_L,\ \mathrm{RR}_U)$.
- The the CI for $\widehat{\mathrm{RRR}}$ is $(1 - \mathrm{RR}_U,\ 1 - \mathrm{RR}_L)$.

Estimating Hazard Ratio

- The Mantel-Haenzsel test calculations enable estimation of the hazard ratio.

$$\begin{aligned}\hat{\psi} &= \frac{O_T/E_T}{O_C/E_C} \\ V[\log\hat{\psi}] &= \frac{1}{E_T} + \frac{1}{E_C}\end{aligned}$$

- Various model-based estimates are preferred where appropriate.

Introduction to Parametric Survival Distributions

- We saw earlier that the Kaplan-Meier estimate of survival is not particularly smooth.
- By assuming a particular parametric form for the survival distribution, smooth estimates of survival can be obtained.
- Some commonly used distributions in survival analysis are:
  - Exponential
  - Weibull
  - Log-normal
  - Log-logistic
- We will look at the exponential and Weibull distributions in more detail.

The Exponential Distribution

- For the exponential distribution with parameter ρ and mean $1/\rho$, we have

$$\mathcal{F}(t) = e^{-\rho t}, f(t) = \rho e^{-\rho t}, h(t) = \rho, H(t) = \rho t$$

- A plot of $H(t) = -\log(\hat{\mathcal{F}}(t))$, where $\hat{\mathcal{F}}(t)$ is the Kaplan-Meier estimate versus t will be linear if the exponential distribution is appropriate.
- The MLE for ρ is $\hat{\rho} = d/\sum x_i$ where d is the total number of failures.

The Exponential Likelihood

- For exponential failure times with d failures we get

$$\begin{aligned} \ell(\rho; x) &= \sum_u \log h(x_i; \phi) + \sum \log \mathcal{F}(x_i; \phi) \\ &= \sum_u \log \rho - \rho \sum x_i \\ &= d \log \rho - \rho \sum x_i \\ U(\rho) &= d/\rho - \sum x_i \\ I(\rho) &= d/\rho^2 \end{aligned}$$

- It is easy to show that solving $U(\rho) = 0$ gives $\hat{\rho} = d/\sum x_i$ as the MLE and $\sqrt{1/I(\hat{\rho})}$ is an estimate of its standard error, although not the best to use in practice.
- Confidence intervals based on the likelihood ratio statistic or from a fitted model are preferred.

Likelihood Ratio Based Confidence Interval

- The likelihood ratio statistic for the null hypothesis $\rho = \rho_0$ is

$$W(\rho_0) = W = 2[\ell(\hat{\rho}) - \ell(\rho_0)]$$

- This has a nice simple form in the exponential case and is distributed approximately as a chi-squared variate with 1 degree of freedom.
- A $1 - \alpha$ confidence region is

$$\{\rho : W(\rho) \leq c^*_{1,\alpha}\}$$

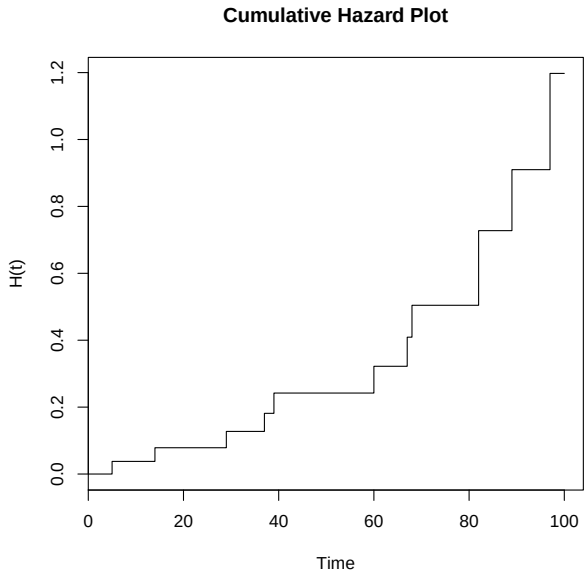
where $c^*_{1,\alpha}$ is the upper α point of the chi-squared distribution with 1 degree of freedom.

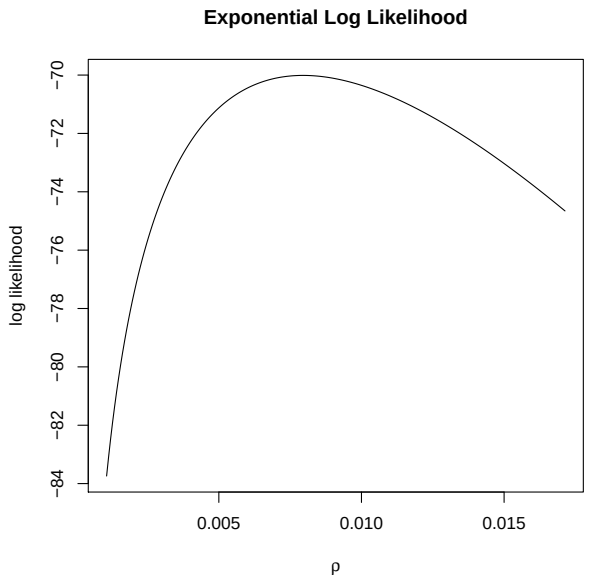
Example

Direct computation of hazard and symmetric 95% CI.

```
> rhohat <- sum(eg1a$status)/sum(eg1a$days)
> se.rhohat <- sqrt(rhohat^2/sum(eg1a$status))
> symci <- rhohat + c(-1, 1) * qnorm(0.975) * se.rhohat
> c(rhohat, symci)
```

```
[1] 0.007952286 0.003452939 0.012451633
```





Example

Model based estimate.

```
> eg1a.sr <- survreg(Surv(days, status) ~ 1, data = eg1a,
+     dist = "exponential")
> summary(eg1a.sr)

Call:
survreg(formula = Surv(days, status) ~ 1, data = eg1a, dist = "exponent
            Value Std. Error    z        p
(Intercept)  4.83      0.289 16.7 6.01e-63

Scale fixed at 1

Exponential distribution
Loglik(model)= -70   Loglik(intercept only)= -70
Number of Newton-Raphson Iterations: 5
n= 29

> exp(-coef(eg1a.sr))

(Intercept)
0.007952286
```

The Two Group Problem

Example

Comparing confidence intervals:

- Symmetric: 0.00345–0.0125
- Likelihood Based: 0.00427–0.0133
- Model Based: 0.00452–0.014

- Suppose you have exponentially distributed failure data for a treatment group T and a control group C.
- Estimate $\hat{\rho}_T = d_T / \sum x_{Ti}$, $\hat{\rho}_C = d_C / \sum x_{Ci}$ and $\hat{\psi} = \hat{\rho}_T / \hat{\rho}_C$.
- It turns out that $V(\log \hat{\psi}) \approx 1/d_T + 1/d_C$ which permits approximate Z-tests and confidence intervals to be calculated.

Example

Direct Calculation

```
> rhohats <- tapply(eg1b$status, eg1b$rx, sum)/tapply(eg1b$days,
+     eg1b$rx, sum)
> hr <- rhohats[2]/rhohats[1]
> lhrse <- sqrt(sum(1/tapply(eg1b$status, eg1b$rx,
+     sum)))
> c(hr, log(hr), lhrse)
```

```
 Treatment  Treatment
 0.6838885 -0.3799604  0.4564355
```

Example

Model based.

```
> eg1b.sr <- survreg(Surv(days, status) ~ rx, data = eg1b,
+     dist = "exponential")
> summary(eg1b.sr)
```

```
Call:
survreg(formula = Surv(days, status) ~ rx, data = eg1b, dist = "exponen
             Value Std. Error      z        p
(Intercept)  4.83      0.289 16.746 6.01e-63
rxTreatment  0.38      0.456  0.832 4.05e-01

Scale fixed at 1

Exponential distribution
Loglik(model)= -119.7   Loglik(intercept only)= -120.1
Chisq= 0.71 on 1 degrees of freedom, p= 0.4
Number of Newton-Raphson Iterations: 5
n= 54
```

Example

Model based (continued).

```
> exp(-coef(eg1b.sr)[2])
```

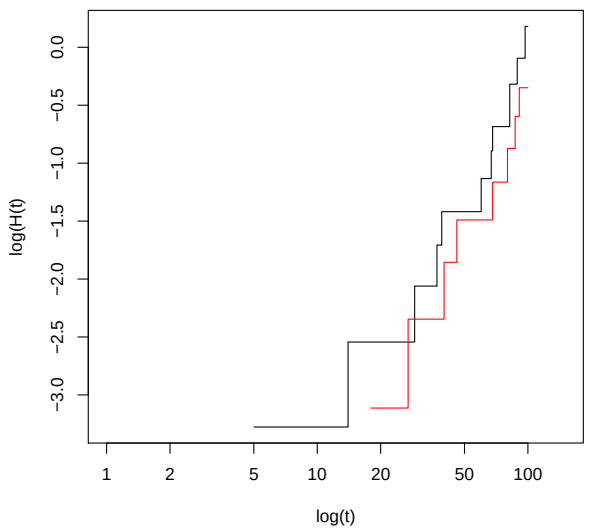
```
rxTreatment
  0.6838885
```

The Weibull Distribution

- The Weibull distribution has the following:

$$\begin{aligned} \mathcal{F}(t) &= \exp[-(\rho t)^{\kappa}] \\ f(t) &= \kappa\rho(\rho t)^{\kappa-1}\exp[-(\rho t)^{\kappa}] \\ h(t) &= \kappa\rho(\rho t)^{\kappa-1} \end{aligned}$$

- Note that if $\kappa = 1$ we are left with an exponential distribution with a hazard of ρ.
- A plot of $\log(-\log(\hat{\mathcal{F}}(t)))$ versus $\log(t)$ will be linear if the Weibull distribution is appropriate.
- The biggest difficulty is that the coefficient estimates don't have nice straight forward interpretations like other models we've used. A scale parameter is estimated and $e^{-\beta_i/\text{scale}}$ gives a hazard ratio.



Example

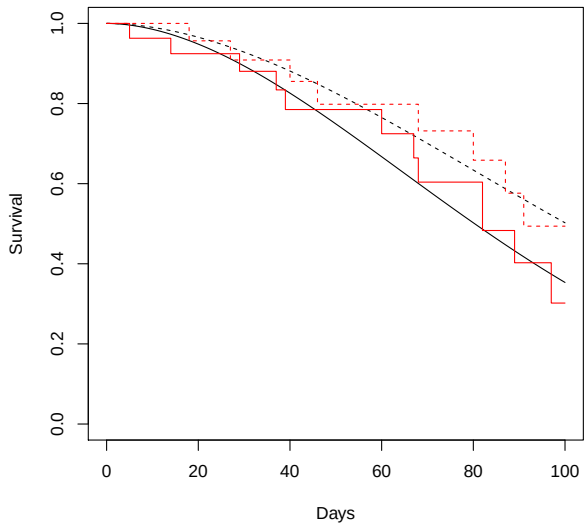
Fitting a Weibull in R

```
> eg2b.sr <- survreg(Surv(days, status) ~ rx, data = eg1b,
+     dist = "weibull")
> summary(eg2b.sr)

Call:
survreg(formula = Surv(days, status) ~ rx, data = eg1b, dist = "weibull
              Value Std. Error      z         p
(Intercept)   4.584      0.162 28.316 2.18e-176
rxTreatment   0.223      0.250  0.894  3.71e-01
Log(scale)   -0.615      0.189 -3.250  1.15e-03

Scale= 0.541

Weibull distribution
Loglik(model)= -115.5   Loglik(intercept only)= -115.9
Chisq= 0.83 on 1 degrees of freedom, p= 0.36
Number of Newton-Raphson Iterations: 8
n= 54
```

Example

The hazard ratio for treatment is given by:

```
> exp(-coef(eg2b.sr)[2]/eg2b.sr$scale)
rxTreatment
  0.6618901
```



Log-logistic and Log-normal Distributions

- These distributions are very similar. Both result in non-proportional hazards models.
- If $\log T$ has a logistic distribution, we say T is *log-logistic*
- If $\log T$ as a normal distribtion, we say T is *log-normal*
- If $\log h(t)$ is non-monotonic in t then these distributions are possible.

Example

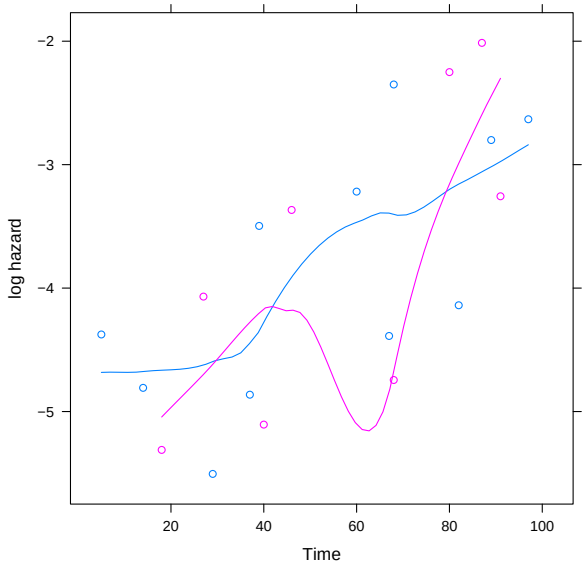
Computing log hazard in R

```
> kmf <- survfit(Surv(days, status) ~ rx, data = eg1b)
> H1 <- -log(kmf[1]$surv)
> H2 <- -log(kmf[2]$surv)
> Ht1 <- eg1b.km[1]$time
> Ht2 <- eg1b.km[2]$time
> dH1 <- diff(H1)/diff(Ht1)
> dH2 <- diff(H2)/diff(Ht2)
> loghaz <- data.frame(ldH = log(c(dH1, dH2)), times = c(Ht1[-1],
+     Ht2[-1]), rx = factor(c(rep(1, length(dH1)),
+     rep(2, length(dH2))), labels = c("Control",
+     "Treatment")))
```

Example

Plotting log hazard in R

```
> library(lattice)
> xyplot(ldH ~ times, data = loghaz, groups = rx,
+     type = c("p", "smooth"), ylab = "log hazard",
+     xlab = "Time")
```



Example

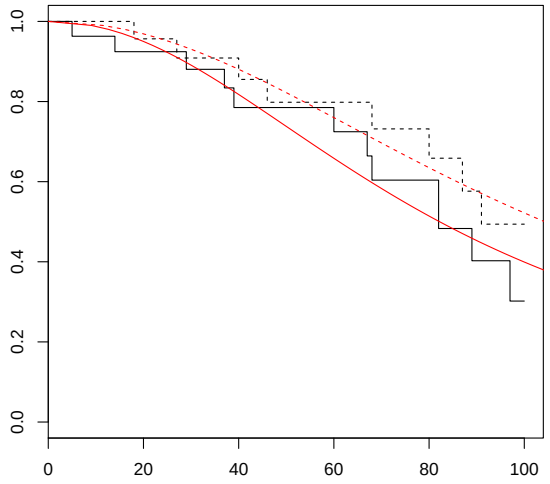
Fitting a log-logistic Model in R

```
> llog.fit <- survreg(Surv(days, status) ~ rx, data = eg1b,
+     dist = "loglogistic")
> summary(llog.fit)

Call:
survreg(formula = Surv(days, status) ~ rx, data = eg1b, dist = "loglogi
             Value Std. Error      z         p
(Intercept)  4.409      0.192 22.969 9.45e-117
rxTreatment  0.238      0.285  0.835  4.04e-01
Log(scale)  -0.734      0.187 -3.931  8.47e-05

Scale= 0.48

Log logistic distribution
Loglik(model)= -116.4   Loglik(intercept only)= -116.8
Chisq= 0.71 on 1 degrees of freedom, p= 0.4
Number of Newton-Raphson Iterations: 5
n= 54
```

Example

Plotting the Result

```
> pct <- 0:99/100
> llog.pred <- predict(llog.fit, newdata = data.frame(rx = factor(1:2,
+     labels = c("Control", "Treatment"))), type = "quantile",
+     p = pct)
> plot(kmf, mark.time = FALSE, lty = 1:2)
> matlines(t(llog.pred), cbind(1 - pct, 1 - pct),
+     col = "red", lty = 1:2)
```

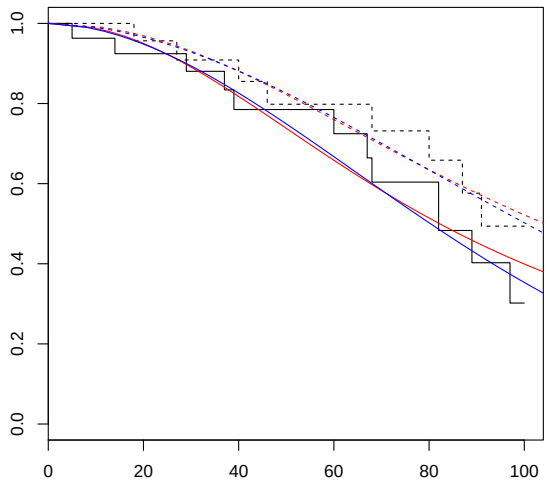


Example

Adding Weibull

```
> weib.pred <- predict(eg2b.sr, , newdata = data.frame(rx = factor(1:2,
+     labels = c("Control", "Treatment"))), type = "quantile",
+     p = pct)
> matlines(t(weib.pred), cbind(1 - pct, 1 - pct),
+     col = "blue", lty = 1:2)
```

References



- Cox D.R. and Oakes D. (1984). *Analysis of Survival Data*. Chapman and Hall
- Kalbfleisch J.D. and Prentice R.L. (1980). *The Statistical Analysis of Failure Time Data*. New York: Wiley