Parametric Models

Steve Simon

The exponential distribution

PDF: $f(t) = \lambda e^{-\lambda t}$ or $f(t) = \frac{1}{\theta} e^{-t/\theta}$

Mean: $E[T] = \frac{1}{\lambda} = \theta$

Think of λ as a failure rate. The faster the failure rate, the shorter the expected failure time.

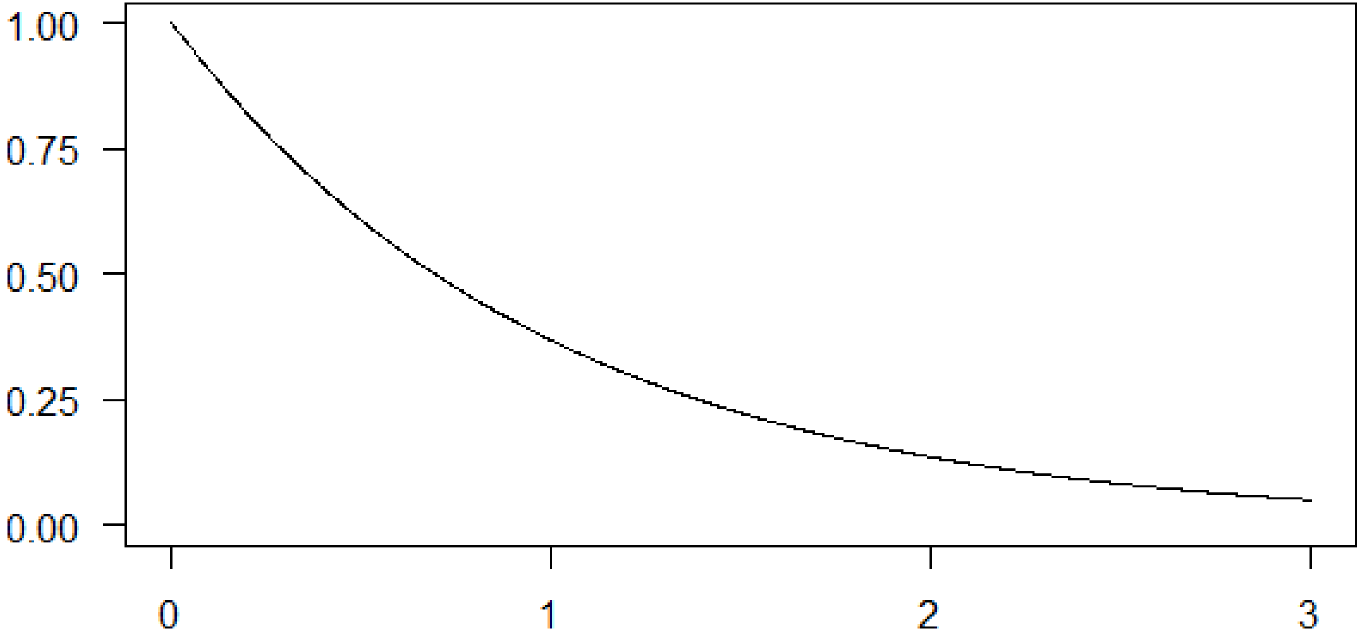
The exponential distribution

CDF: $F(t) = 1 - e^{-\lambda t} = 1 - e^{t/\theta}$

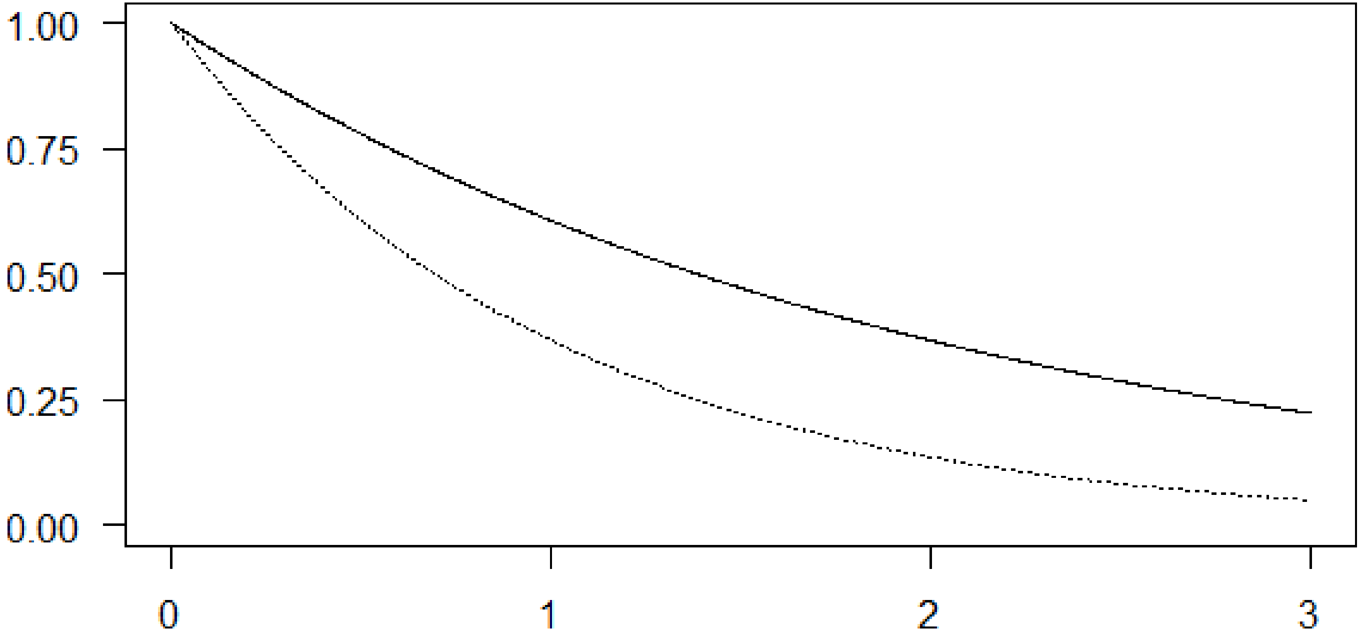
Survival: $S(t) = e^{-\lambda t} = e^{-t/\theta}$

Hazard: $h(t) = \lambda = \frac{1}{\theta}$

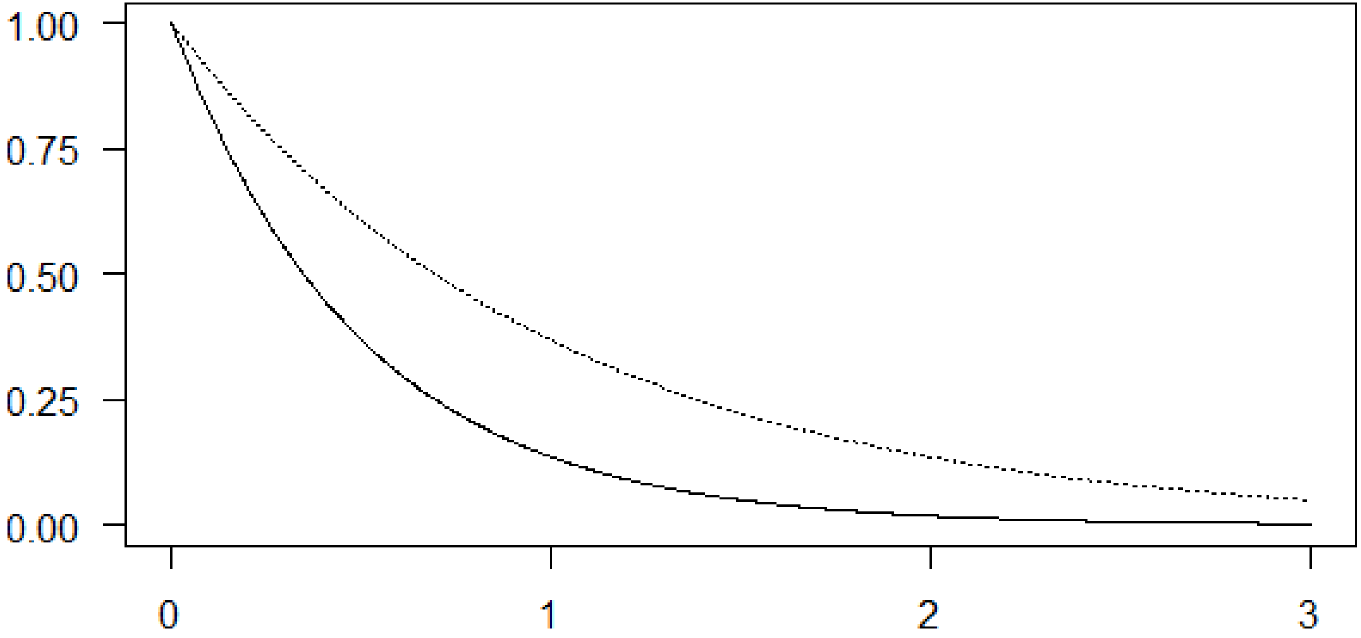
Exponential survival curve for theta=1



Exponential survival curve for theta=2



Exponential survival curve for theta=0.5



The Weibull distribution

Survival: $S(t) = e^{-(t/\theta)^k}$

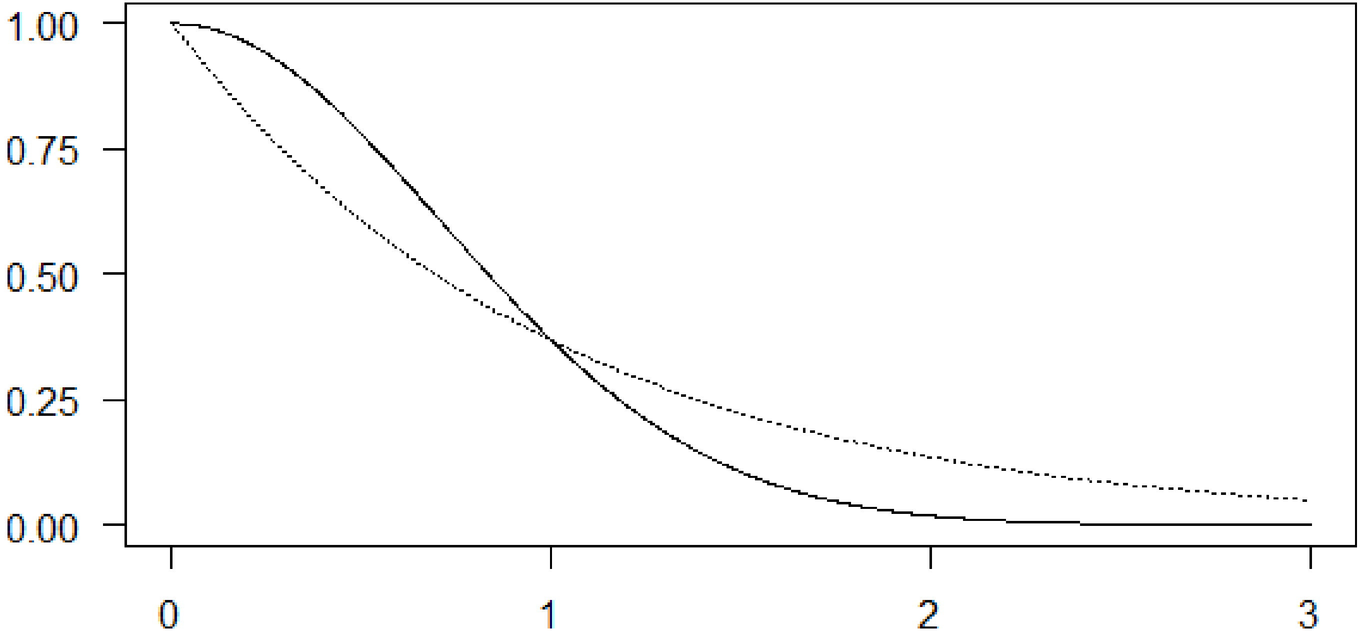
Hazard: $h(t) = \frac{k}{\theta}\left(\frac{t}{\theta}\right)^{k-1}$

PDF: $f(t) = \frac{k}{\theta}\left(\frac{t}{\theta}\right)^{k-1} e^{-(t/\theta)^k}$ or $f(t) = \frac{1}{\theta} e^{-t/\theta}$

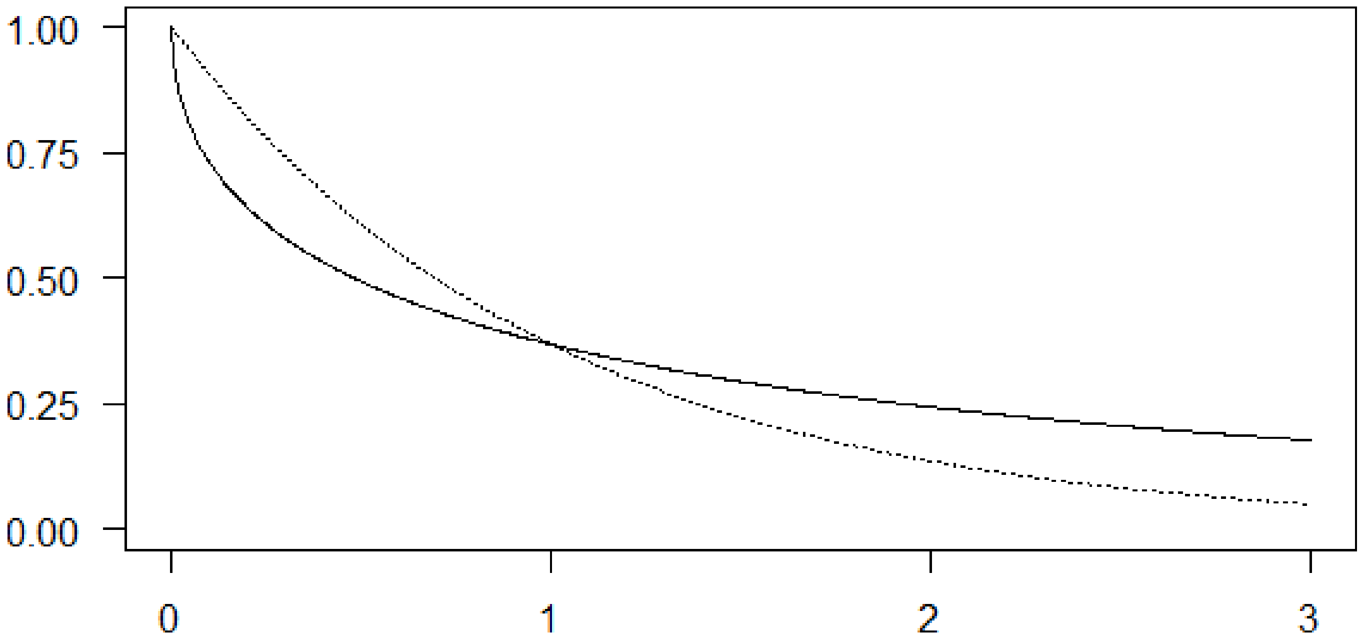
Relationship between k and hazard for the Weibull distribution

- $k > 1$, increasing hazard rate.
- $k < 1$, decreasing hazard rate.
- $k = 1$, constant hazard rate (exponential distribution)

Weibull survival curve for theta=2



Weibull survival curve for k=0.5



How Wikipedia presents the Weibull distribution, part 1

Standard parameterization [edit]

The probability density function of a Weibull random variable is:[1]

$$f(x;\lambda,k)=\begin{cases}\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}e^{-(x/\lambda)^{k}} & x\geq 0,\\ 0 & x<0,\end{cases}$$

where $k > 0$ is the *shape parameter* and $\lambda > 0$ is the *scale parameter* of the distribution. Its complementary cumulative distribution function is a stretched exponential function.

Excerpt from Wikipedia page

How Wikipedia presents the Weibull distribution, part 2

Alternative parameterizations [edit]

Applications in medical statistics and econometrics often adopt a different parameterization.[4][5] The shape parameter k is the same as above, while the scale parameter is $b = \lambda^{-k}$. In this case, for $x \geq 0$, the probability density function is

$$f(x; k, b) = bkx^{k-1}e^{-bx^k},$$

Excerpt from Wikipedia page

How Wikipedia presents the Weibull distribution, part 3

A third parameterization can also be found.[6][7] The shape parameter *k* is the same as in the standard case, while the scale parameter is $\beta = 1/\lambda$. Then, for *x* ≥ 0, the probability density function is

$$f(x; k, \beta) = \beta k(\beta x)^{k-1} e^{-(\beta x)^k}$$

the cumulative distribution function is

$$F(x; k, \beta) = 1 - e^{-(\beta x)^k},$$

and the hazard function is

$$h(x; k, \beta) = \beta k(\beta x)^{k-1}.$$

Excerpt from Wikipedia page

A different presentation of the Weibull distribution

The Weibull density with shape parameter λ and scale parameter θ is

$$f_Y(y) = \frac{\lambda y^{\lambda-1}}{\theta^\lambda} \exp\left[-\left(\frac{y}{\theta}\right)^\lambda\right].$$

The survivor function is

$$\begin{aligned} S_Y(y) &= \int_y^\infty \frac{\lambda t^{\lambda-1}}{\theta^\lambda} \exp\left[-\left(\frac{t}{\theta}\right)^\lambda\right] dt \\ &= \exp\left[-\left(\frac{y}{\theta}\right)^\lambda\right]. \end{aligned}$$

The hazard function is

$$\begin{aligned} h_Y(y) &= \frac{f_Y(y)}{S_Y(y)} \\ &= \frac{\frac{\lambda y^{\lambda-1}}{\theta^\lambda} \exp\left[-\left(\frac{y}{\theta}\right)^\lambda\right]}{\exp\left[-\left(\frac{y}{\theta}\right)^\lambda\right]} \\ &= \left(\frac{\lambda}{\theta^\lambda}\right) y^{\lambda-1}. \end{aligned}$$

Excerpt from people.stat.sfu.ca/~raltman/stat402/402L32.pdf

The accelerated time model (exponential distribution)

Recall that the survival function for the exponential distribution is

Survival: $S(t) = e^{-t/\theta}$

The accelerated time model replaces θ with

$$e^{(\beta_0+\beta_1 X)}$$

This produces the survival curve

$$S(t, X, \beta_0, \beta_1) = e^{-t/e^{(\beta_0+\beta_1 X)}}$$

The values of β_0 and β_1 will end up stretching or shrinking the time scale.

Percentiles

The pth percentile of the accelerated time model is

$$-ln(1-p)e^{\beta_0+\beta_1 X}$$

and the ratio of two percentiles, one with X=X_1 and the other with X=X_2 is

$$\frac{-ln(1-p)e^{\beta_0+\beta_1 X_1}}{-ln(1-p)e^{\beta_0+\beta_1 X_2}} = e^{\beta_1(X_1-X_2)}$$

If X_1 is one unit larger than X_2, this reduces to e^{β_1}.

Different, but not different

This may look quite different than the model we used for Kaplan-Meier curves and the Cox proportional hazards model, but it actually is not. The hazard function is

$$h(t, x, \beta_0, \beta_1) = e^{-(\beta_0 + \beta_1 x)}.$$

Notice that the hazard is constant with respect to t. The baseline hazard, the hazard when X=0 is

$$h_0(t) = e^{-\beta_0}$$

Different, but not different

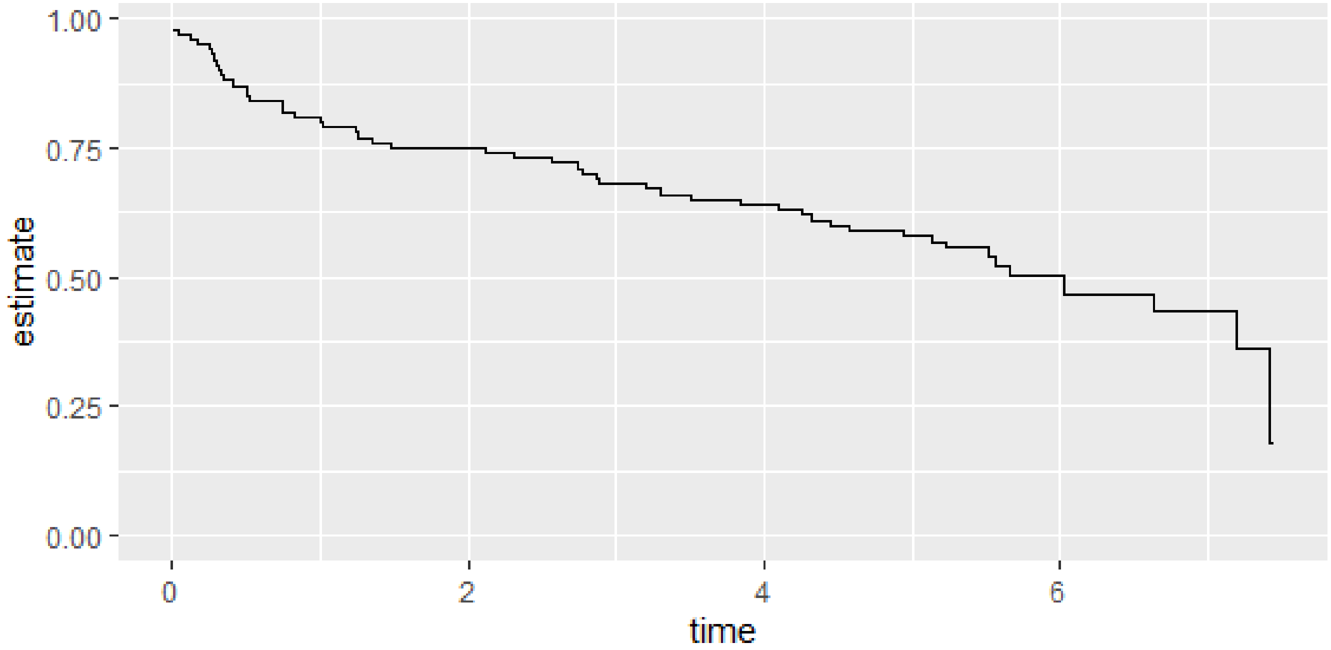
The hazard ratio for a subject with $x = x_1$ compared to a subject with $x = x_2$ is

$$e^{-\beta_1(x_1-x_2)}$$

Example with whas100 dataset

```
##   id  admitdate    foldate los lenfol fstat
## 1  1 03/13/1995 03/19/1995   4      6  Dead
## 2  2 01/14/1995 01/23/1996   5    374  Dead
## 3  3 02/17/1995 10/04/2001   5   2421  Dead
##   age gender      bmi  time_yrs age_group
## 1  65   Male 31.38134 0.0164271     60-69
## 2  88 Female 22.65790 1.0239562      >=80
## 3  77   Male 27.87892 6.6283368     70-79
```

Before you fit the exponential models, fit a Kaplan-Meier curve.



Estimate an exponential model

```
##
## Call:
## survreg(formula = whas100_surv ~ 1, dist = "exponential")
##              Value Std. Error    z      p
## (Intercept)  2.09       0.14 14.9 <2e-16
##
## Scale fixed at 1
##
## Exponential distribution
## Loglik(model)= -157.6   Loglik(intercept only)= -157.6
## Number of Newton-Raphson Iterations: 5
## n= 100
```

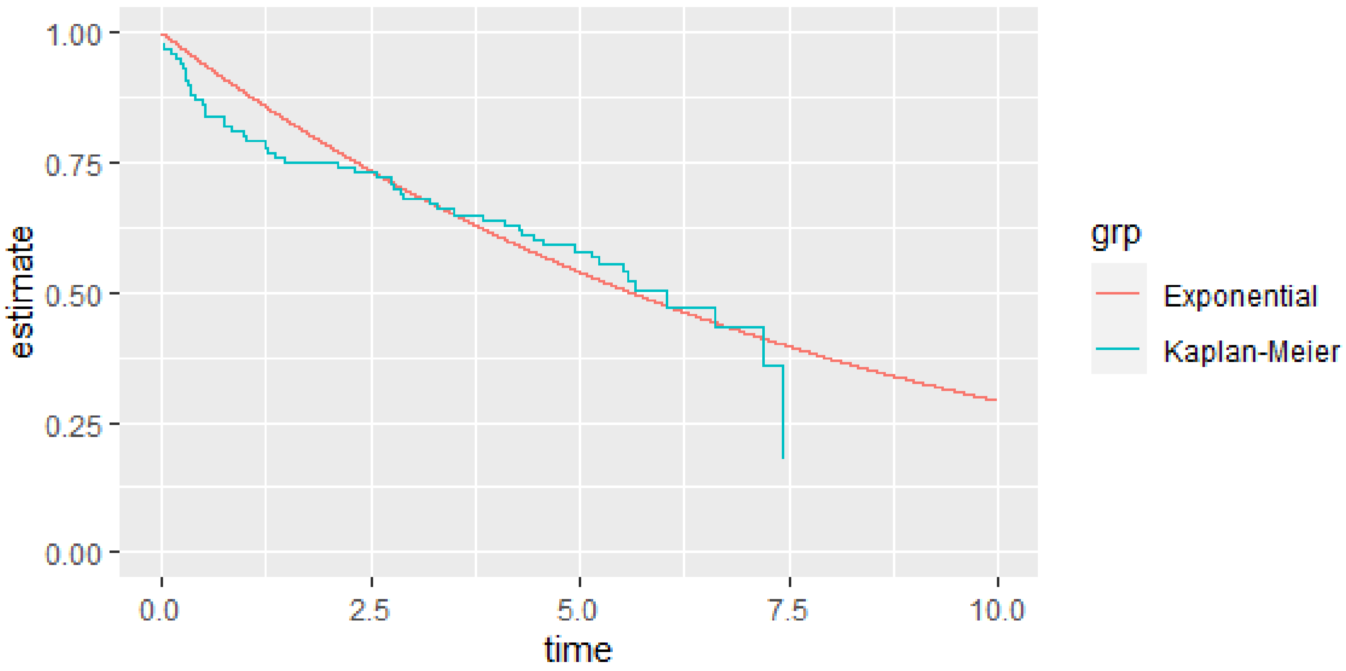
Percentile calculations

75th percentile: -ln(0.75) exp(2.09)=2.3

50th percentile: -ln(0.5) exp(2.09)=5.6

25th percentile: -ln(0.25) exp(2.09)=11.2

Comparison of exponential fit to Kaplan-Meier



Weibull model

Recall that the survival function for a Weibull distribution is

$$S(t) = e^{-(t/\theta)^k}$$

Replace θ with

$$e^{(\beta_0+\beta_1 X)}$$

to get

$$S(t) = e^{-(t/e^{(\beta_0+\beta_1 X)})^k}$$

Alternate Weibull formulation

The original development of parametric survival models chose a different, but equivalent formulation.

$$ln(T) = \beta_0 + \beta_1 X + \sigma\varepsilon$$

where ε has a log-Weibull distribution (also known as a Gompertz or extreme value distribution).

The parameter σ is equal to $1/k$.

Estimate a Weibull model

```
##
## Call:
## survreg(formula = whas100_surv ~ 1, dist = "weibull")
##               Value Std. Error     z       p
## (Intercept) 2.268      0.212 10.71 <2e-16
## Log(scale)  0.312      0.129  2.41  0.016
##
## Scale= 1.37
##
## Weibull distribution
## Loglik(model)= -154.3   Loglik(intercept only)= -154.3
## Number of Newton-Raphson Iterations: 5
## n= 100
```

SAS table of Weibull fit

The LIFEREG Procedure

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	2.2676	0.2117	1.8528	2.6825	114.79	<.0001
Scale	1	1.3655	0.1767	1.0596	1.7596		
Weibull Scale	1	9.6566	2.0438	6.3777	14.6211		
Weibull Shape	1	0.7323	0.0947	0.5683	0.9437		

What are those two extra values in the SAS code?

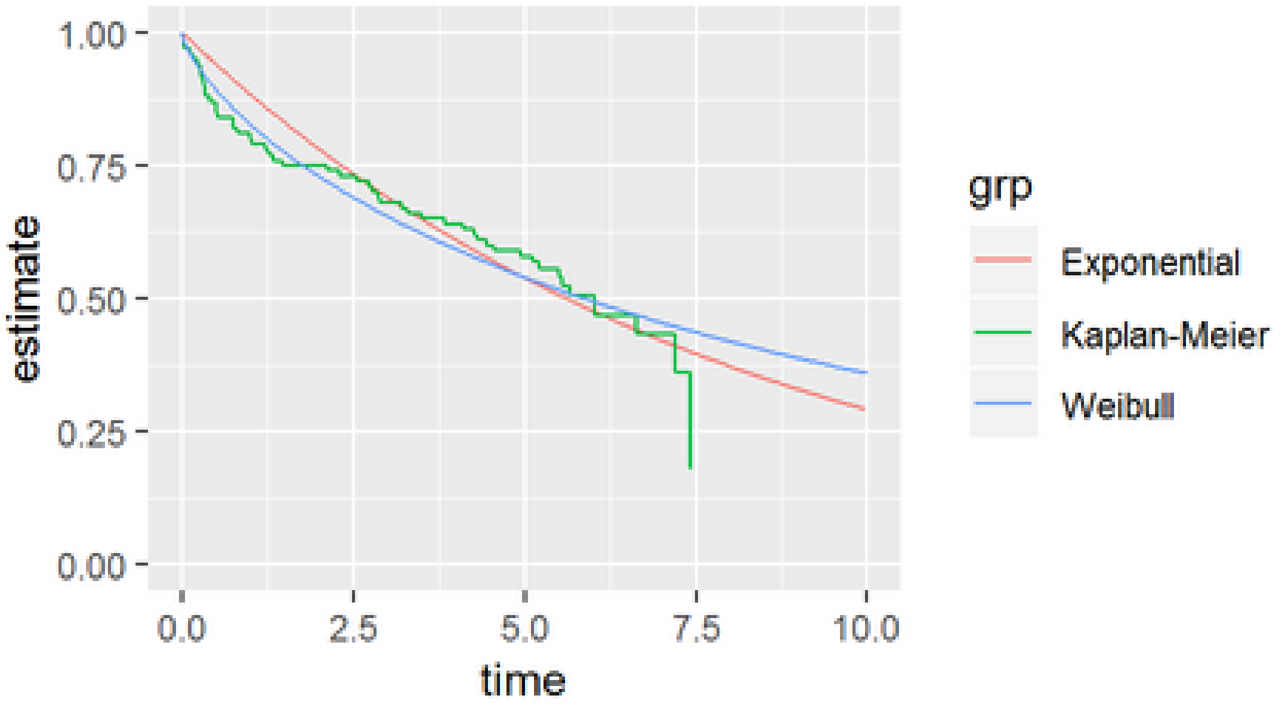
The LIFEREG Procedure

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	2.2676	0.2117	1.8528	2.6825	114.79	<.0001
Scale	1	1.3655	0.1767	1.0596	1.7596		
Weibull Scale	1	9.6566	2.0438	6.3777	14.6211		
Weibull Shape	1	0.7323	0.0947	0.5683	0.9437		

exp(2.2676)

1 / 1.3655

Graph of Weibull survival



Estimate gender effect in a Weibull model

```
##           term   estimate std.error statistic
## 1  (Intercept)  2.5631446 0.2844093  9.012168
## 2 genderFemale -0.7904360 0.3914775 -2.019110
## 3   Log(scale)  0.3033633 0.1284480  2.361760
##        p.value  conf.low   conf.high
## 1 2.020223e-19  2.005713  3.12057667
## 2 4.347583e-02 -1.557718 -0.02315417
## 3 1.818842e-02        NA          NA
```

Cluster and frailty models

A fundamental assumption of all the models so far is the assumption of independence.

- Whether the event time for one patient is early or late has no effect on the event time for a different patient.

There are settings, however, where two event times are correlated, and you can model this correlation using a cluster effect or a frailty effect.

Examples of correlated survival times

- Times to failure of pairs of organs (kidneys, eyes) within a patient
- Multi-center trials
- Recurrent events (infection, re-hospitalization)

Sandwich estimate

You can account for cluster effects by modifying the variance-covariance of the coefficients. For the normal model without frailty effects, the information matrix is

$$I(\beta) = \frac{\partial^2 L_p}{(\partial\beta)^2}$$

and the estimated variance covariance matrix, $\hat{V}$, is

$$\hat{V} = I(\hat{\beta})^{-1}$$

Sandwich

The robust sandwich estimator (similar to the sandwich estimator in Generalized Estimating Equations) is

$$\hat{R} = \hat{V}(\hat{L}'\hat{L})\hat{V}$$

where $\hat{L}$ is the vector of Schoenfeld residuals. The middle of this sandwich, the $\hat{L}'\hat{L}$, adjusts the variance covariance matrix to account for the correlation within clusters.

Frailty model

Recall that the proportional hazards model assumes that the hazard function of a given patient,

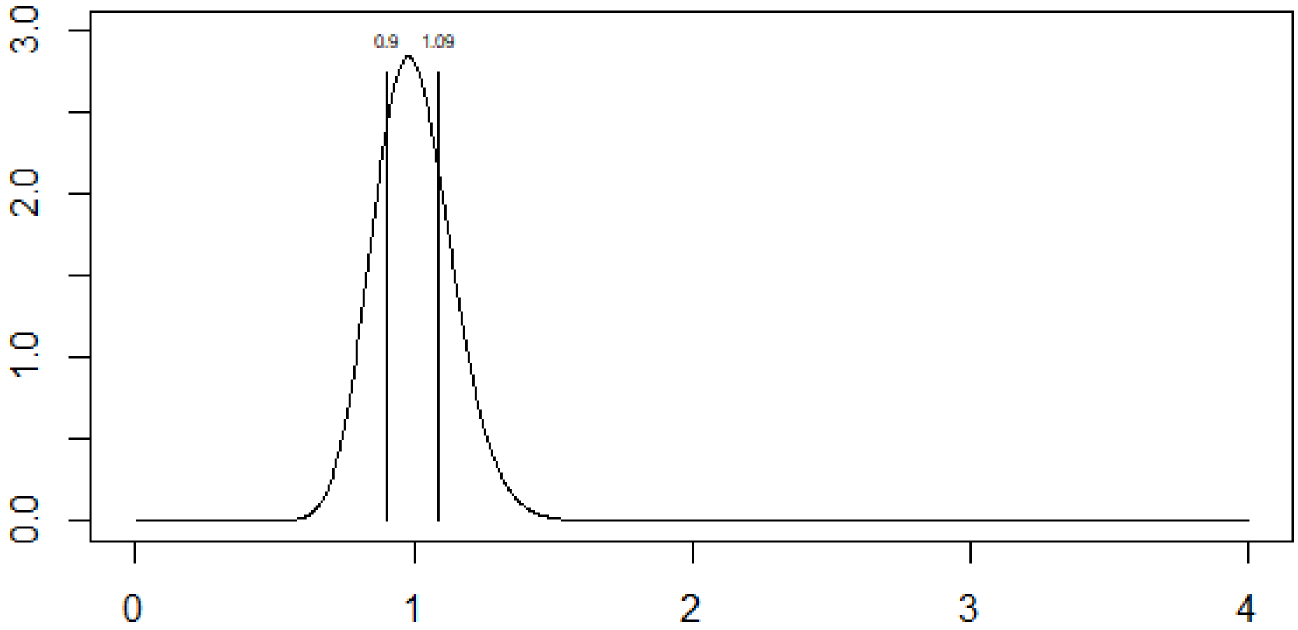
$$h(t, X, \beta) = h_0(t)e^{X\beta}$$

The frailty model multiplies all the hazards within a family f by a frailty term z_f.

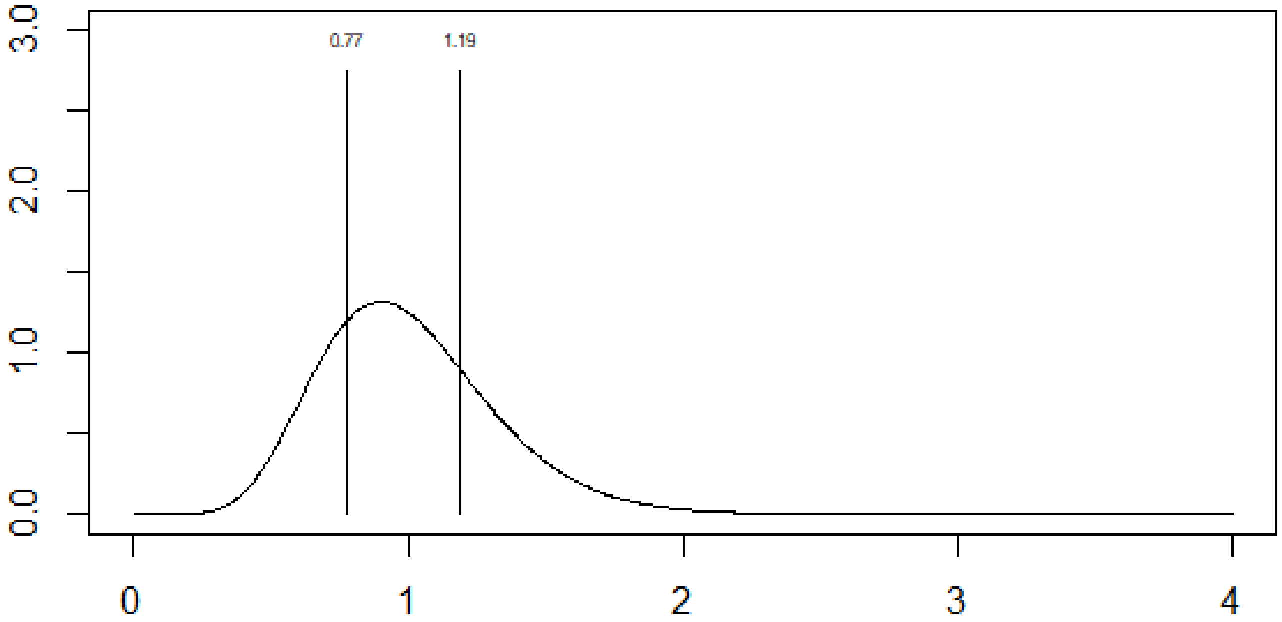
$$h(t, X, \beta) = z_f h_0(t)e^{X\beta}$$

Typically, z_f is given a gamma distribution with a mean of 1 and a variance of $\frac{1}{\alpha}$.

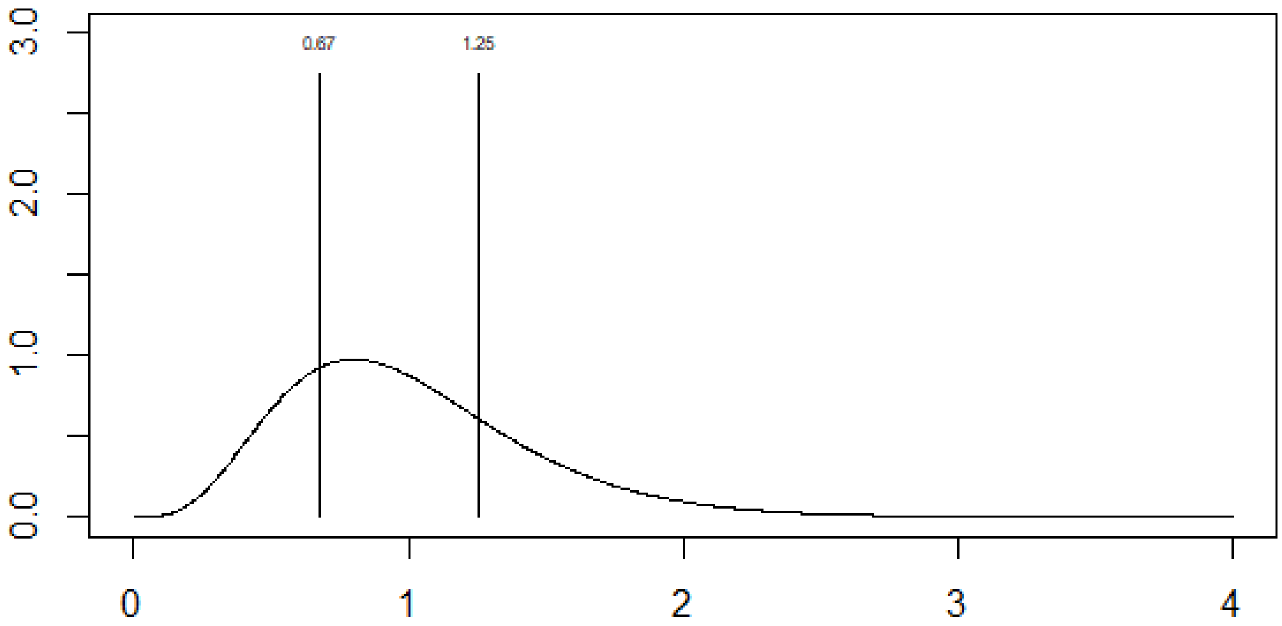
Gamma distribution, alpha=50



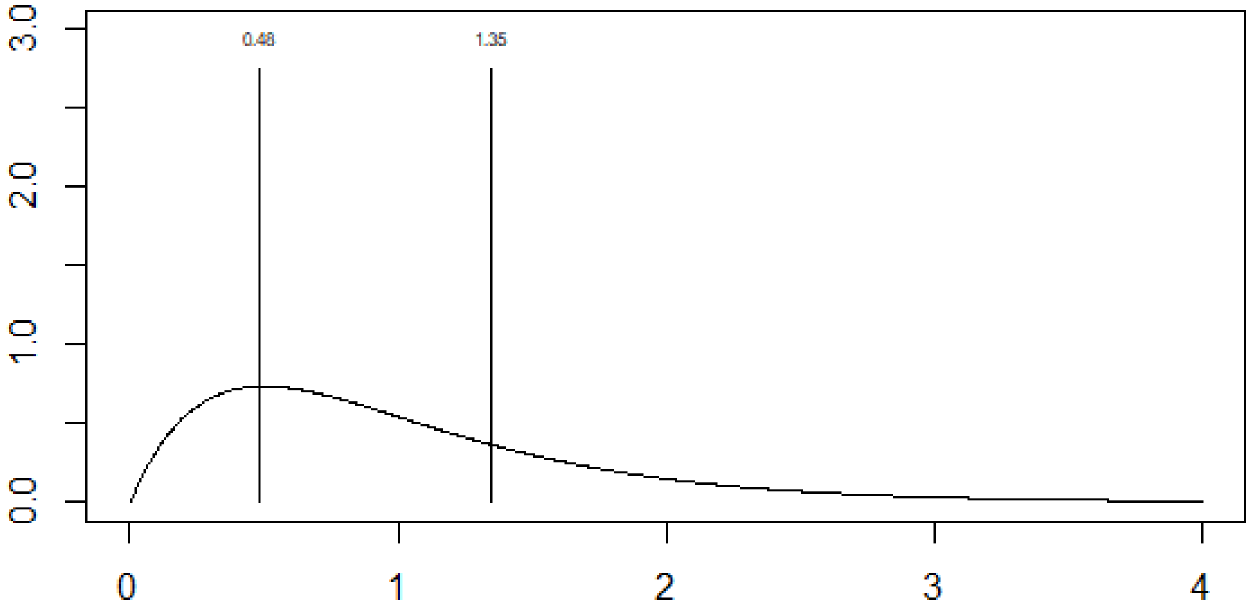
Gamma distribution, alpha=10



Gamma distribution: alpha=5



Gamma distribution: alpha=2



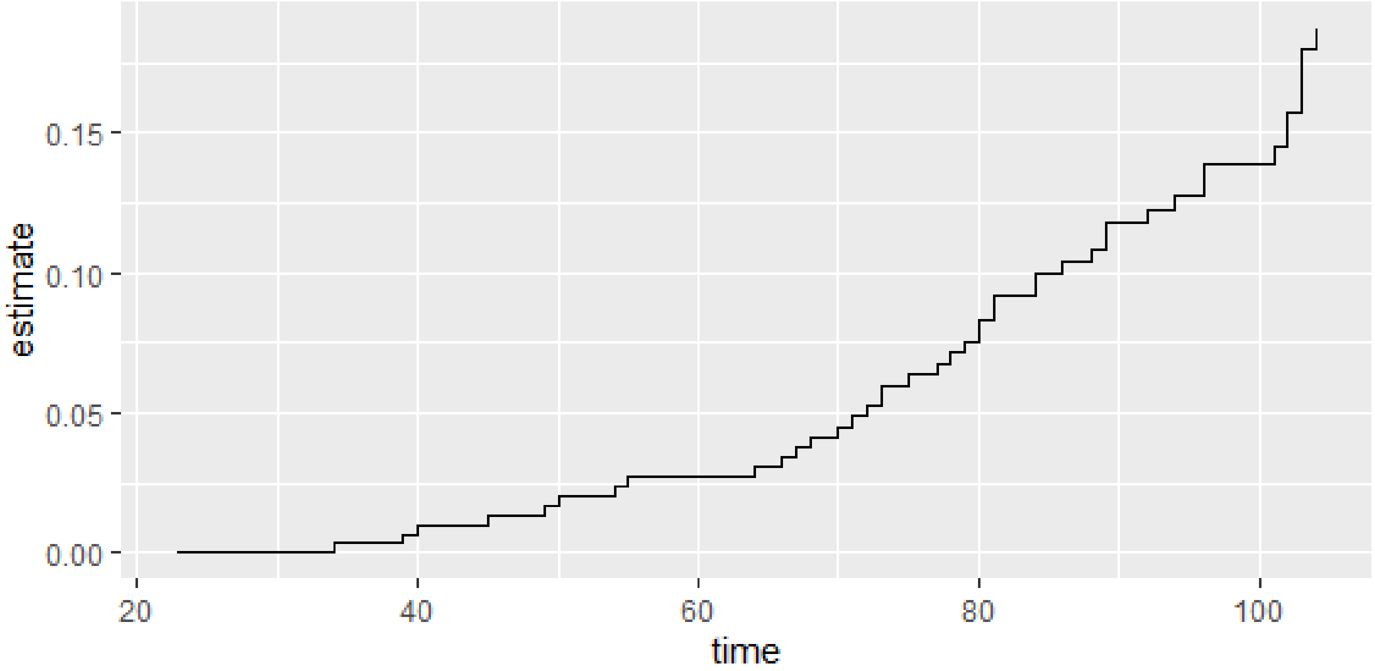
Partial listing of rats data

```
##   litter rx time status sex
## 1      1  1  101      0   f
## 2      1  0   49      1   f
## 3      1  0  104      0   f
## 4      2  1   91      0   m
## 5      2  0  104      0   m
## 6      2  0  102      0   m
```

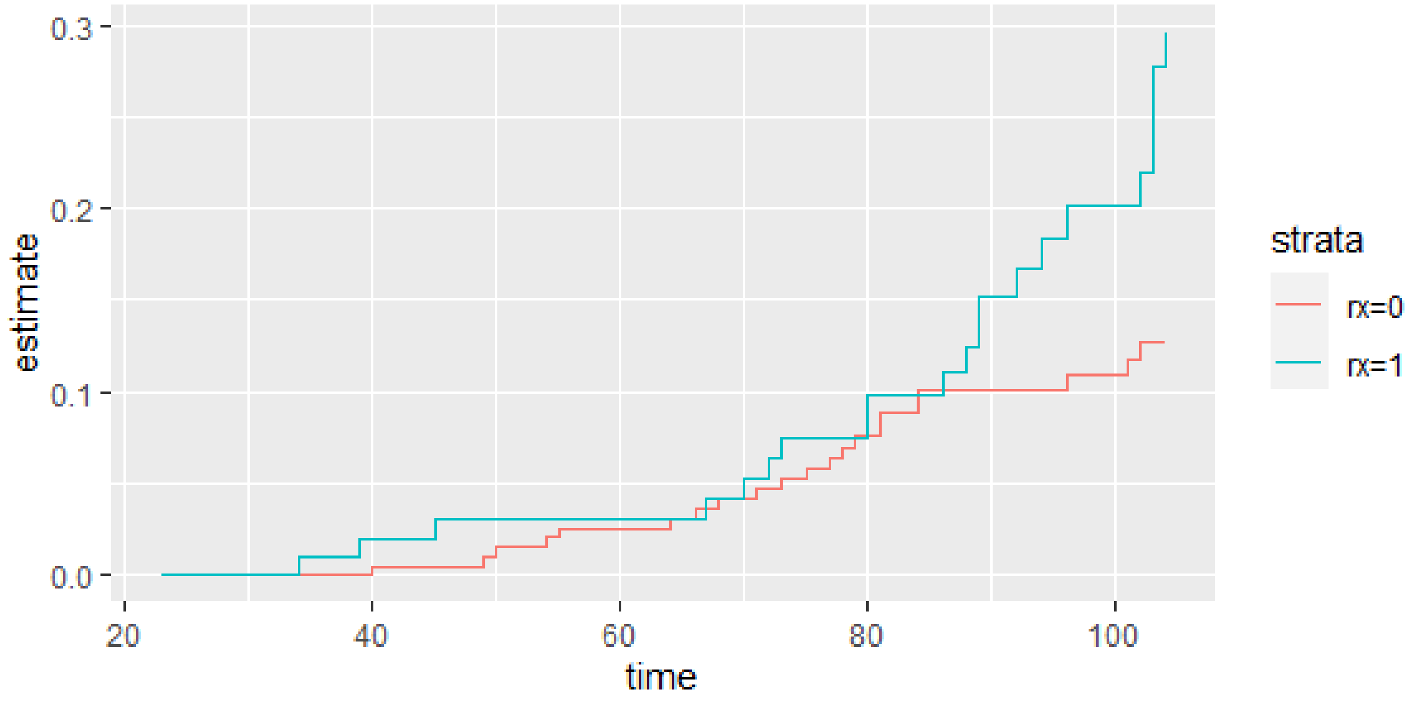
Descriptive statistics for rats data

```
##
##   0   1
## 200 100
##
##   0   1
## 258  42
##
##   f   m
## 150 150
##    Min. 1st Qu.  Median    Mean 3rd Qu.
##   23.00   80.75   98.00   90.44  104.00
##    Max.
##  104.00
```

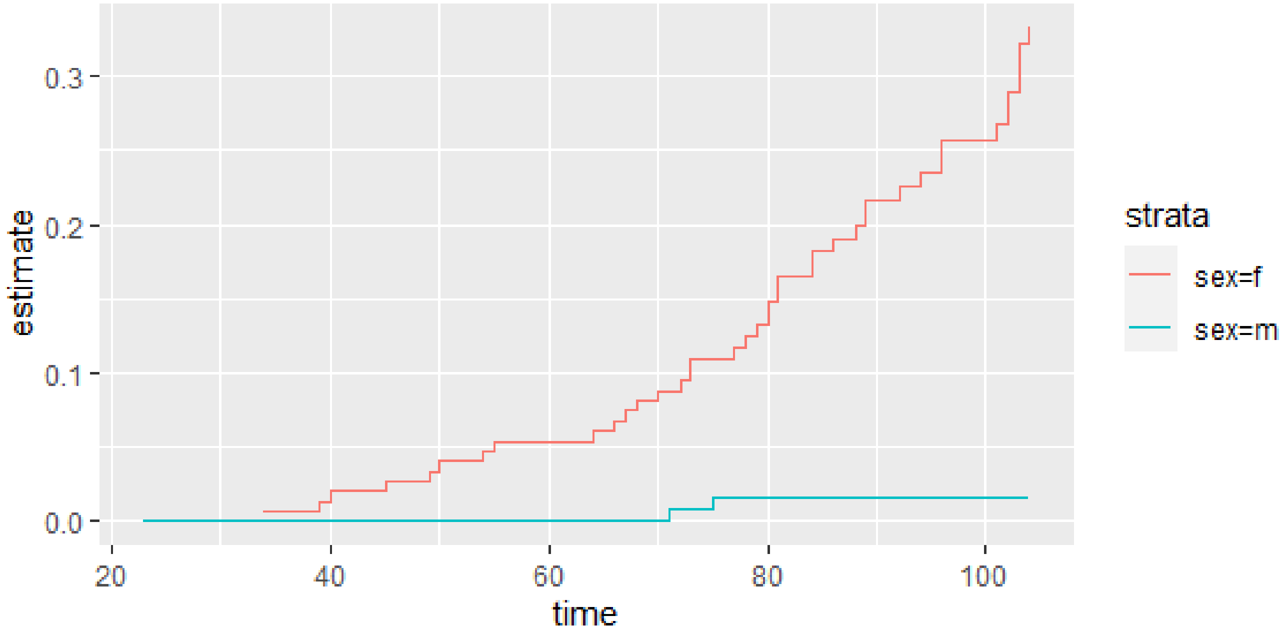
Overall survival



Survival by treatment



Survival by sex



Fit a model ignoring the effect of litter

```
## Call:
## coxph(formula = rats_surv ~ rx, data = rats, subset = (sex ==
##     "f"))
##
##       coef exp(coef) se(coef)     z       p
## rx 0.9047    2.4713   0.3175 2.849 0.00438
##
## Likelihood ratio test=7.98  on 1 df, p=0.004741
## n= 150, number of events= 40
```

Cluster effect

```
Call:
coxph(formula = rats_surv ~ rx + cluster(litter), data = rats,
    subset = (sex == "f"))

    coef exp(coef) se(coef) robust se    z      p
rx 0.905     2.471    0.318     0.303 2.99 0.0028

Likelihood ratio test=7.98  on 1 df, p=0.00474
n= 150, number of events= 40
```

Frailty effect

```
Call:
coxph(formula = rats_surv ~ rx + frailty(litter), data = rats,
    subset = (sex == "f"))

                  coef se(coef)   se2  Chisq   DF      p
rx               0.914    0.323 0.319  8.012  1.0 0.0046
frailty(litter)                       17.692 14.4 0.2443

Iterations: 6 outer, 24 Newton-Raphson
     Variance of random effect= 0.499   I-likelihood = -180.8
Degrees of freedom for terms=  1.0 14.4
Likelihood ratio test=37.6  on 15.4 df, p=0.00124  n= 150
```