Survival Analysis
Session 2: From Data to Distribution

Jonas Schöley

@jschoeley
0000-0002-3340-8518
j.schoeley@uni-rostock.de

Recap: Survival Identities

In survival analysis we consider the random variable
“*X*: Time until event”
x = 0.1 weeks, 2.3 weeks…

Recap: Survival Identities

In survival analysis we consider the random variable
“*X*: Time until event”
x = 0.1 weeks, 2.3 weeks…

We express our knowledge about the distribution of *X* in any of these functions. Knowing any single function we can derive all other via the **Survival Identities**.

Recap: Survival Identities

In survival analysis we consider the random variable **"*X*: Time until event"**
x = 0.1 weeks, 2.3 weeks…

We express our knowledge about the distribution of *X* in any of these functions. Knowing any single function we can derive all other via the **Survival Identities**.

f(*x*): Density function
The relative likelihood of experiencing the event around time *x*.

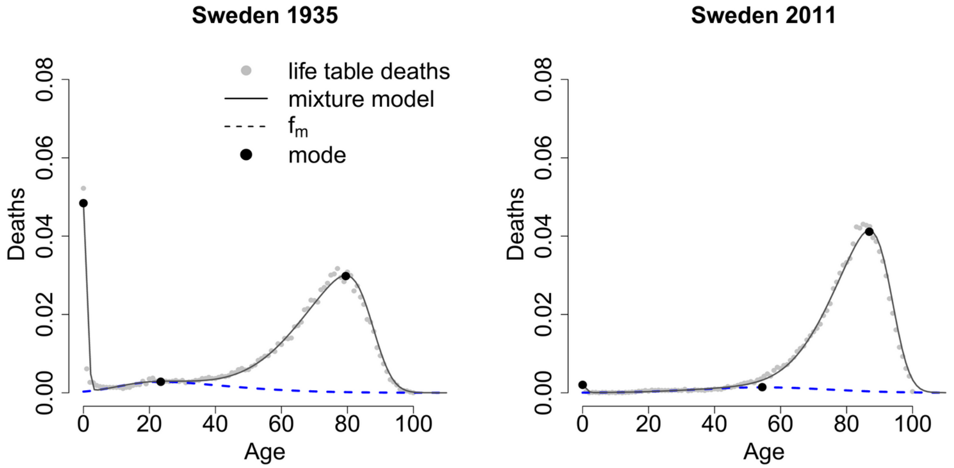


Fig. 2 Model fit on life table deaths for Sweden in 1935 and 2011. The solid line shows the overall mixture model. The dotted line highlights the fit of the Skew Normal employed to estimate accidental and premature mortality. The big dots point out the three modal ages of the distribution

Zanotto etal. (2021). [A Mixture-Function Mortality Model](#).

Recap: Survival Identities

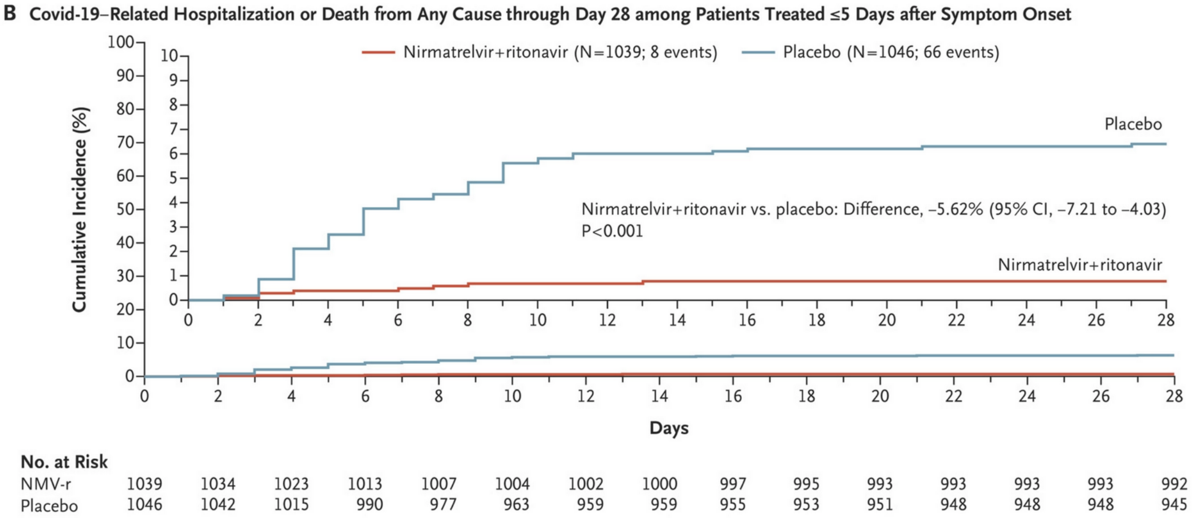
In survival analysis we consider the random variable
"*X*: Time until event"
x = 0.1 weeks, 2.3 weeks…

We express our knowledge about the distribution of *X* in any of these functions. Knowing any single function we can derive all other via the **Survival Identities**.

f(*x*): Density function
The relative likelihood of experiencing the event around time *x*.

$F(x) = \int_0^x f(x)\,dx$

F(*x*): Distribution function
aka Cumulative function
The probability of experiencing the event until time *x*.
$F(x) = P(X \leq x)$



Hammond etal. (2022).
Oral Nirmatrelvir for High-Risk, Nonhospitalized Adults with Covid-19.

Recap: Survival Identities

In survival analysis we consider the random variable
"*X*: Time until event"
x = 0.1 weeks, 2.3 weeks…

We express our knowledge about the distribution of *X* in any of these functions. Knowing any single function we can derive all other via the **Survival Identities**.

f(*x*): Density function
The relative likelihood of experiencing the event around time *x*.

$F(x)={}_0\int^x f(x)\,dx$

$S(x)={}_x\int^{\mathrm{Inf}} f(x)\,dx$

F(*x*): Distribution function
aka Cumulative function
The probability of experiencing the event until time *x*.
$F(x) = P(X \leq x)$

$S(x)=1-F(x)$

S(*x*): Survival function

The probability of *not* experiencing the event until time *x*.
$S(x) = P(X>x)$

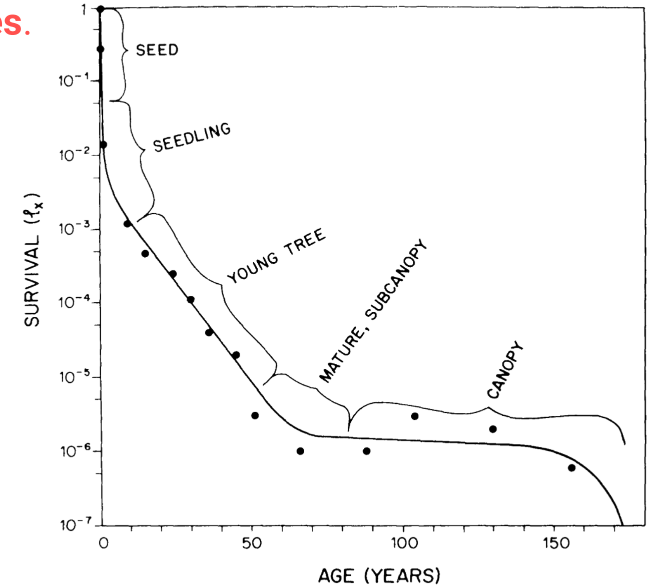


FIGURE 1. Survivorship curve for *Euterpe globosa*, over 7 orders of magnitude. All values of l_x except that at 9 years are completely independent of each other, thus providing an internal check on accuracy.

Valen (1975). [Life, Death, and Energy of a Tree.](#)

Recap: Survival Identities

In survival analysis we consider the random variable
“*X*: Time until event”
x = 0.1 weeks, 2.3 weeks…

We express our knowledge about the distribution of *X* in any of these functions. Knowing any single function we can derive all other via the **Survival Identities**.

h(*x*): Hazard function
The instantaneous rate of new events at time *x* among those who did not experience the event yet.
$h(x) = \lim_{h\to 0} P(x \leq X < x+h \mid X \geq x)/h$

f(*x*): Density function
The relative likelihood of experiencing the event around time *x*.

$F(x)={}_0\int^x f(x)\,dx$

$S(x)={}_x\int^{\mathrm{Inf}} f(x)\,dx$

$h(x)=-S'(x)/S(x)$

F(*x*): Distribution function
aka Cumulative function
The probability of experiencing the event until time *x*.
$F(x) = P(X \leq x)$

$S(x)=1-F(x)$

S(*x*): Survival function

The probability of *not* experiencing the event until time *x*.
$S(x) = P(X > x)$

Recap: Survival Identities

In survival analysis we consider the random variable **"*X*: Time un**
x = 0.1 week

We express our knowledge about the distribution of *X* in
function we
es.

Klüver (2022). The survival of interest groups.

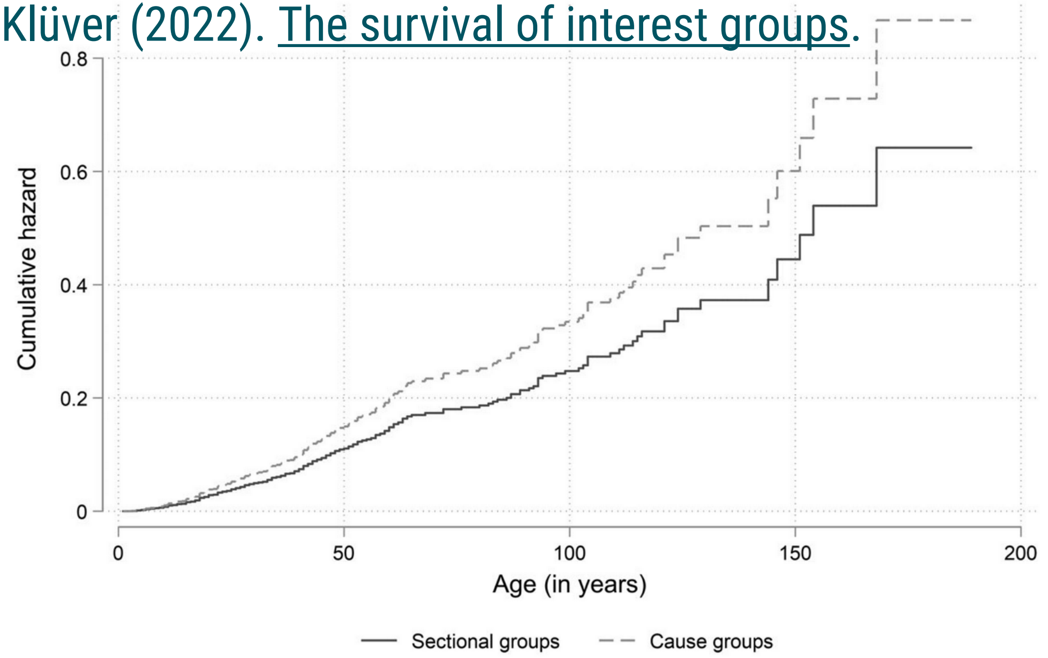


Figure 3. The effect of interest group type.

h(*x*): Hazard function
The instantaneous rate of new events at time *x* among those who did not experience the event yet.
$h(x) = \lim_{h\to 0} P(x \leq X < x+h | X \geq x)/h$

$F(x) = {}_0\int^x$

$H(x) = {}_0\int^x h(x)\,dx$

$= -\log S(x)$

H(*x*): Cumulative Hazard
The integral of h(x).

F(*x*): Distributi
aka Cumulative
The probability
experiencing th
time *x*.
$F(x) = P(X \leq x)$

time *x*.
$S(x) = P(X > x)$

$S(x) = \exp(-H(x))$

Last Week's Homework

Choose a time-to-event setting that interests you and look up a constant rate related to that setting. What is the time scale for your setting? When does the time-to-event start? When have have half of the population experienced the event given the chosen rate?

Example: Today we looked at the time until I catch COVID. I choose the rate 2,428 infections per 100,000 persons per 7 days from the local COVID incidences and assumed this rate to be constant. The timescale was "weeks into the semester" and it starts at the first week of the semester. I used the survival function of the exponential distribution to calculate the time until the probability of catching COVID reached 50%.
$S(x) = \exp(-{}_0\!\int^x \lambda dx) = \exp(-\lambda x)$

What Does Survival Data Look Like?

In survival analysis we consider the random variable **“*X*: Time until event”**
x = 0.1 month, 2.3 months…

What Does Survival Data Look Like?

In survival analysis we consider the random variable
"*X*: Time until event"
x = 0.1 month, 2.3 months…

Who? 8 breast cancer patients

What Does Survival Data Look Like?

In survival analysis we consider the random variable
“*X*: Time until event”
x = 0.1 month, 2.3 months…

Who? 8 breast cancer patients
Start of observation? Cancer diagnosis

What Does Survival Data Look Like?

In survival analysis we consider the random variable **"*X*: Time until event"**
x = 0.1 month, 2.3 months…

Who? 8 breast cancer patients
Start of observation? Cancer diagnosis
Event of interest? Death

What Does Survival Data Look Like?

In survival analysis we consider the random variable
“*X*: Time until event”
x = 0.1 month, 2.3 months…

Who? 8 breast cancer patients
Start of observation? Cancer diagnosis
Event of interest? Death
Time unit? Months

What Does Survival Data Look Like?

In survival analysis we consider the random variable **“*X*: Time until event”**
x = 0.1 month, 2.3 months…

Who? 8 breast cancer patients
Start of observation? Cancer diagnosis
Event of interest? Death
Time unit? Months
End of observation? 10 years follow up

What Does Survival Data Look Like?

In survival analysis we consider the random variable **"*X*: Time until event"**
x = 0.1 month, 2.3 months…

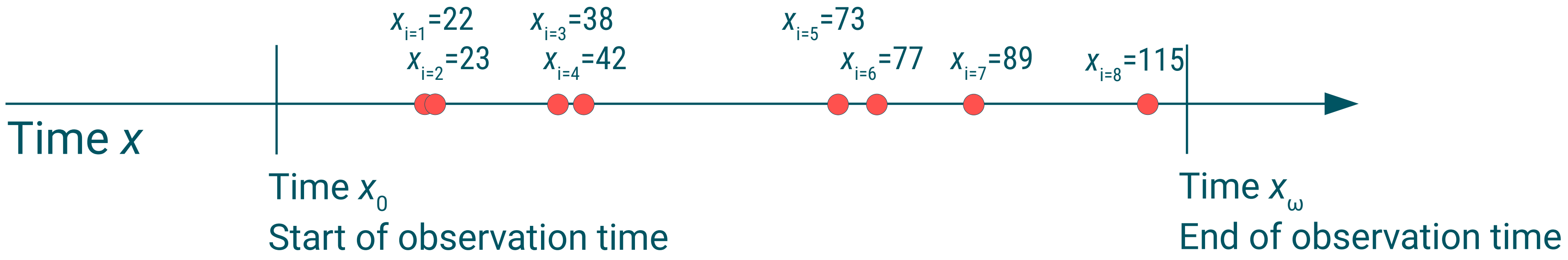
Who? 8 breast cancer patients
Start of observation? Cancer diagnosis
Event of interest? Death
Time unit? Months
End of observation? 10 years follow up

What Does Survival Data Look Like?

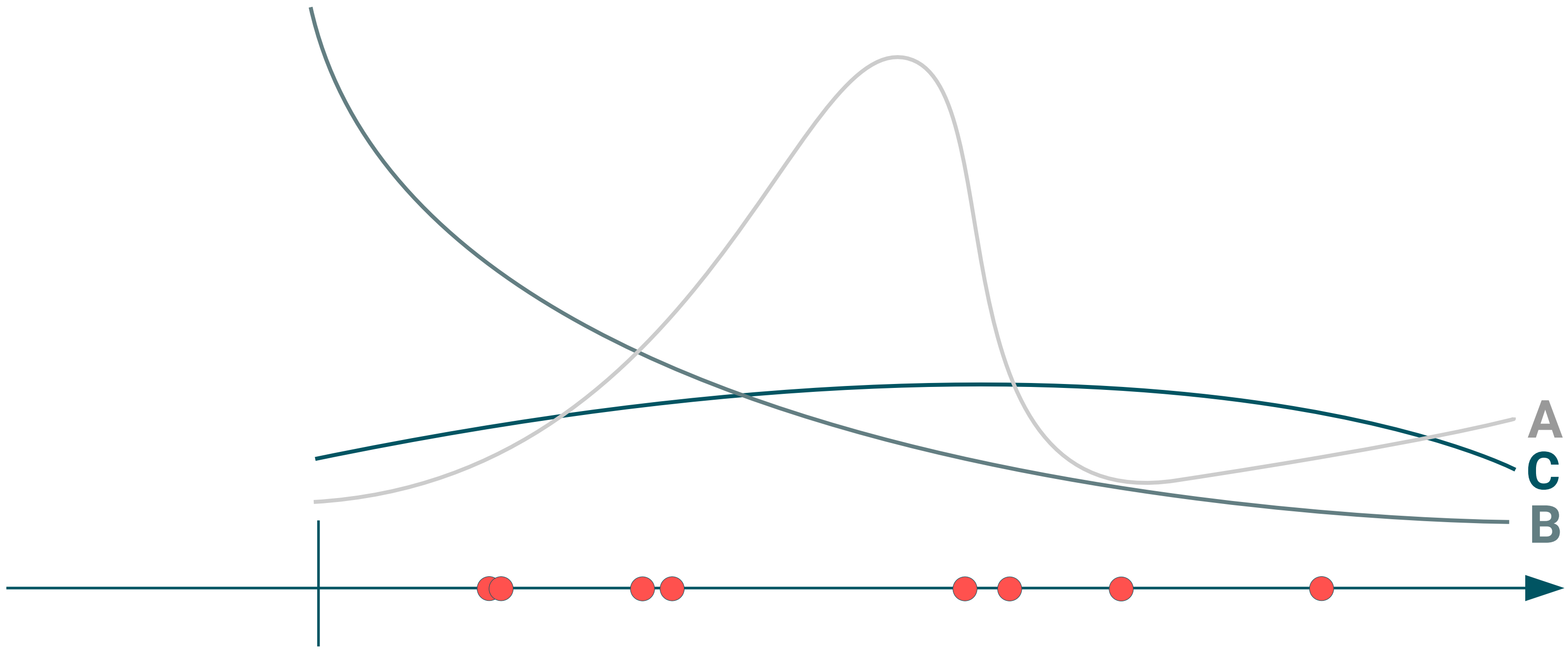
In survival analysis we consider the random variable
“*X*: Time until event”
x = 0.1 month, 2.3 months…

Who? 8 breast cancer patients
Start of observation? Cancer diagnosis
Event of interest? Death
Time unit? Months
End of observation? 10 years follow up

Observation index *i*	Observed event time *x*
1	22
2	23
3	38
4	42
5	73
6	77
7	89
8	115

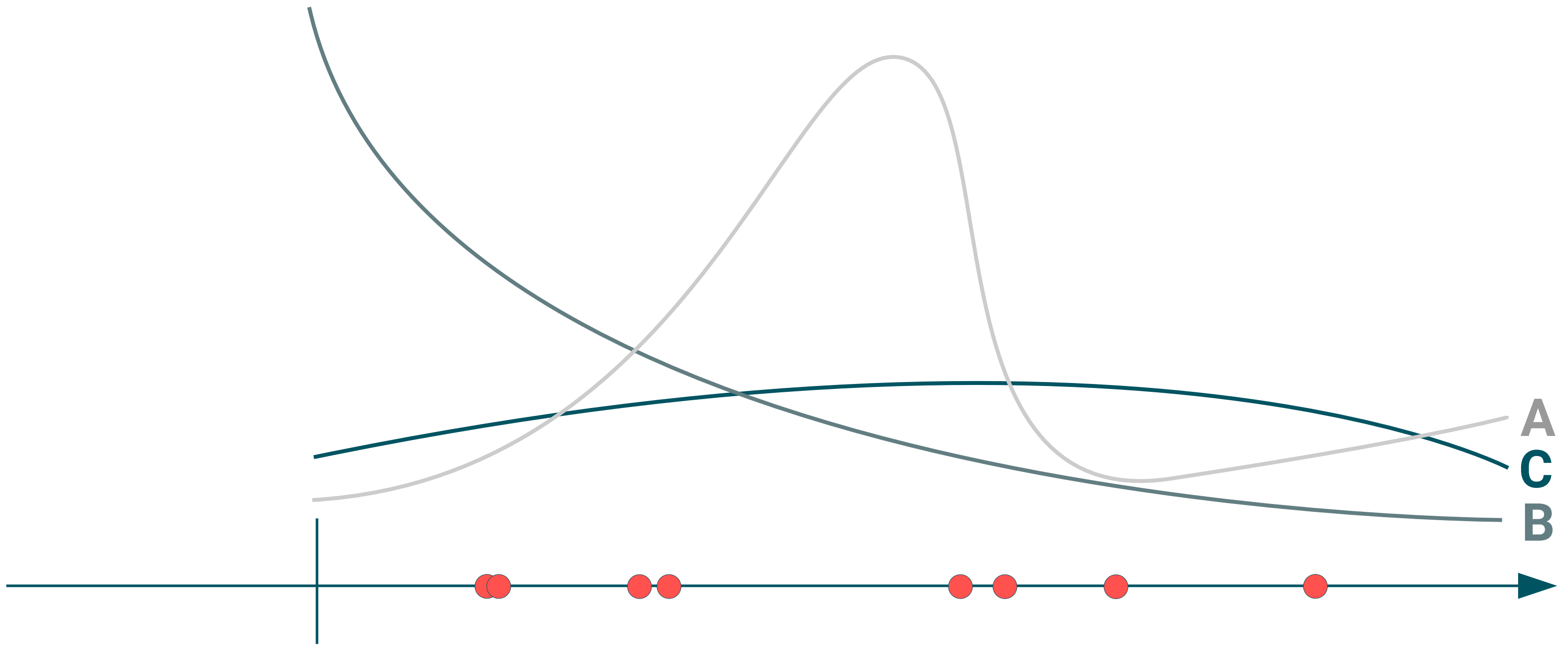


Inferring the Survival Density from Data



Which distribution most likely corresponds to the data?

Inferring the Survival Density from Data



Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Maximum Likelihood Estimation

We fit a model $f_{\boldsymbol{\theta}}$ to the data $\boldsymbol{x}$ by choosing model parameters $\boldsymbol{\theta}$ which maximize the **likelihood function** L, i.e. which make the observed data most probable.

Product over all observations

Probability density given parameters $\boldsymbol{\theta}$

$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_i f_{\boldsymbol{\theta}}(x_i)$

Single observed survival time

In practice we often maximize the **log-likelihood** for convenience:

$\log L(\boldsymbol{\theta}|\boldsymbol{x}) = \log \prod_i f_{\boldsymbol{\theta}}(x_i) = \sum_i \log f_{\boldsymbol{\theta}}(x_i)$

Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Maximum Likelihood Estimation

We fit a model $f_{\boldsymbol{\theta}}$ to the data $\boldsymbol{x}$ by choosing model parameters $\boldsymbol{\theta}$ which maximize the **likelihood function** L, i.e. which make the observed data most probable.

Product over all observations

Probability density given parameters $\boldsymbol{\theta}$

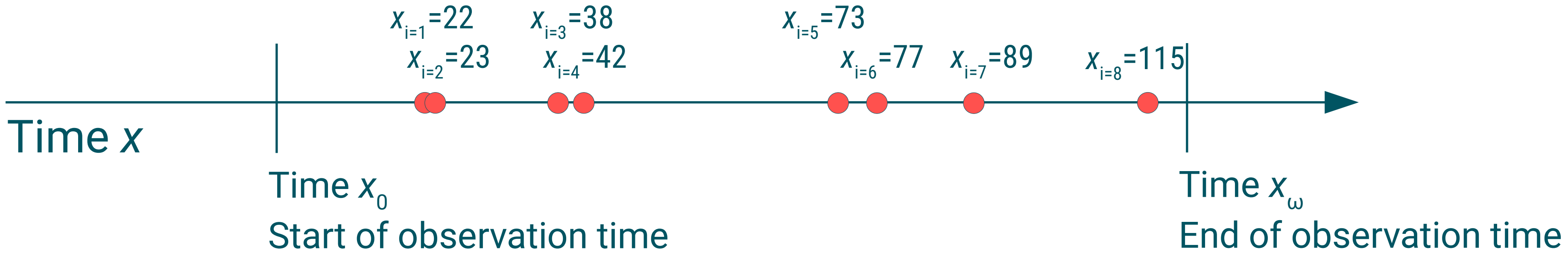
$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_i f_{\boldsymbol{\theta}}(x_i)$ Single observed survival time

In practice we often maximize the **log-likelihood** for convenience:

$\log L(\boldsymbol{\theta}|\boldsymbol{x}) = \log \prod_i f_{\boldsymbol{\theta}}(x_i) = \sum_i \log f_{\boldsymbol{\theta}}(x_i)$

You need:

→ data $\boldsymbol{x}$



Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Maximum Likelihood Estimation

We fit a model $f_{\boldsymbol{\theta}}$ to the data $\boldsymbol{x}$ by choosing model parameters $\boldsymbol{\theta}$ which maximize the **likelihood function** L, i.e. which make the observed data most probable.

Product over all observations

Probability density given parameters $\boldsymbol{\theta}$

$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_i f_{\boldsymbol{\theta}}(x_i)$ Single observed survival time

In practice we often maximize the **log-likelihood** for convenience:

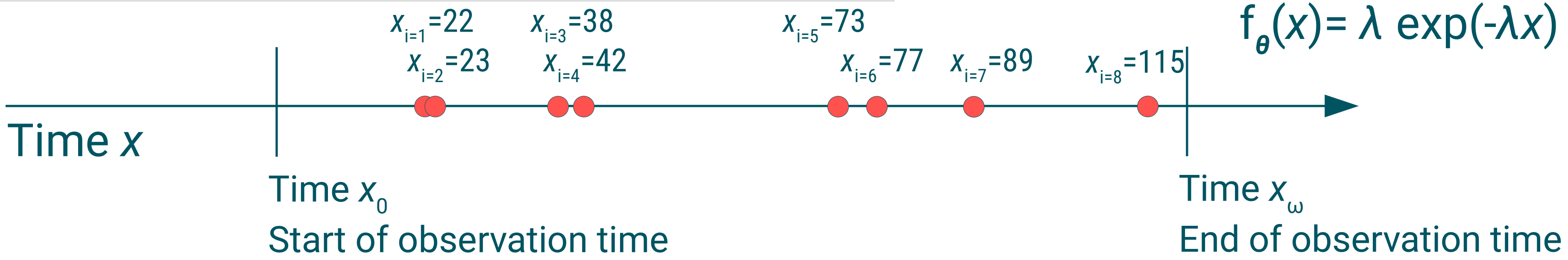
$\log L(\boldsymbol{\theta}|\boldsymbol{x}) = \log \prod_i f_{\boldsymbol{\theta}}(x_i) = \sum_i \log f_{\boldsymbol{\theta}}(x_i)$

You need:

→ data $\boldsymbol{x}$

→ a probability density $f_{\boldsymbol{\theta}}$ ("the model") parameterized by…

$f_{\boldsymbol{\theta}}(x) = \lambda \exp(-\lambda x)$



Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Maximum Likelihood Estimation

We fit a model $f_{\boldsymbol{\theta}}$ to the data $\boldsymbol{x}$ by choosing model parameters $\boldsymbol{\theta}$ which maximize the **likelihood function** L, i.e. which make the observed data most probable.

Product over all observations

Probability density given parameters $\boldsymbol{\theta}$

$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_i f_{\boldsymbol{\theta}}(x_i)$

Single observed survival time

In practice we often maximize the **log-likelihood** for convenience:

$\log L(\boldsymbol{\theta}|\boldsymbol{x}) = \log \prod_i f_{\boldsymbol{\theta}}(x_i) = \sum_i \log f_{\boldsymbol{\theta}}(x_i)$

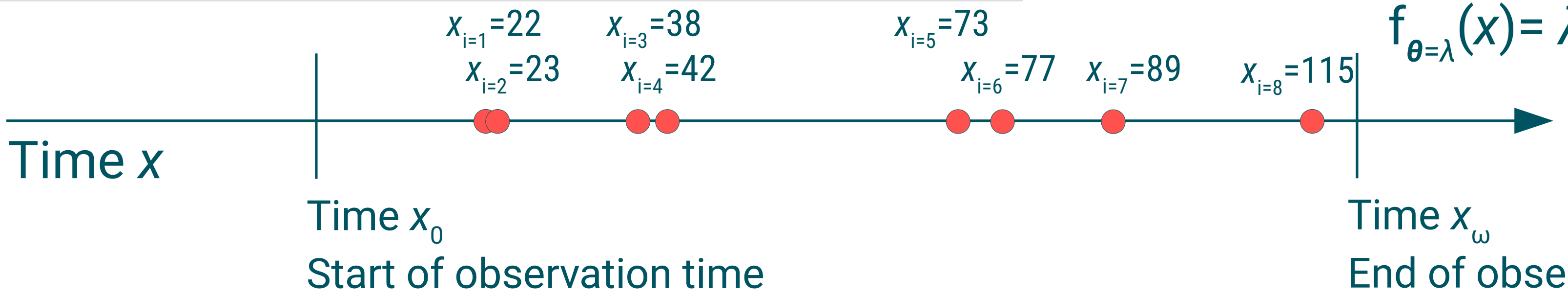
You need:

→ data $\boldsymbol{x}$

→ a probability density $f_{\boldsymbol{\theta}}$ (“the model”) parameterized by…

→ …a set of parameters $\boldsymbol{\theta}$

$f_{\boldsymbol{\theta}=\lambda}(x) = \lambda \exp(-\lambda x)$

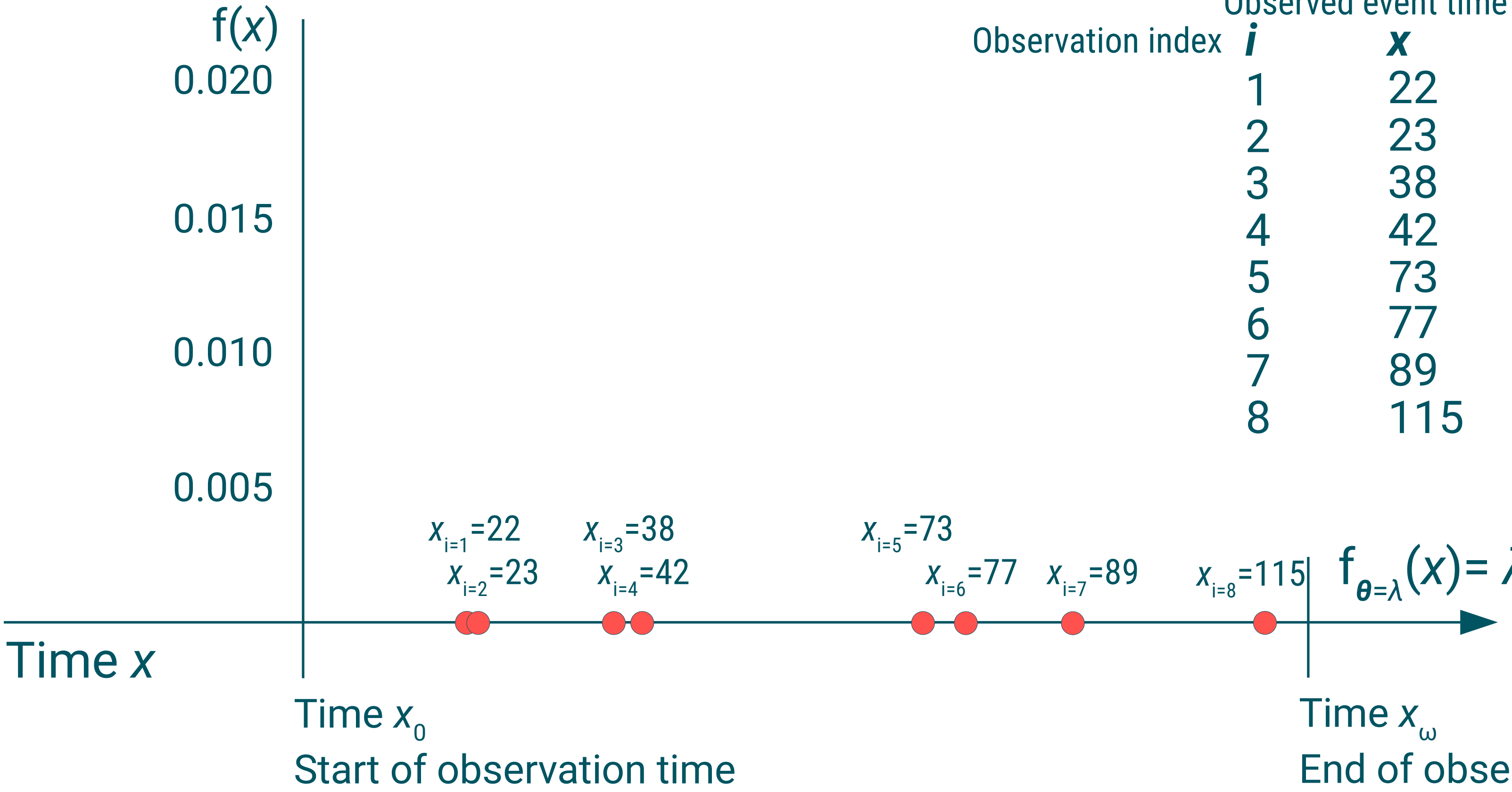


Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Observation index i	Observed event time x
1	22
2	23
3	38
4	42
5	73
6	77
7	89
8	115



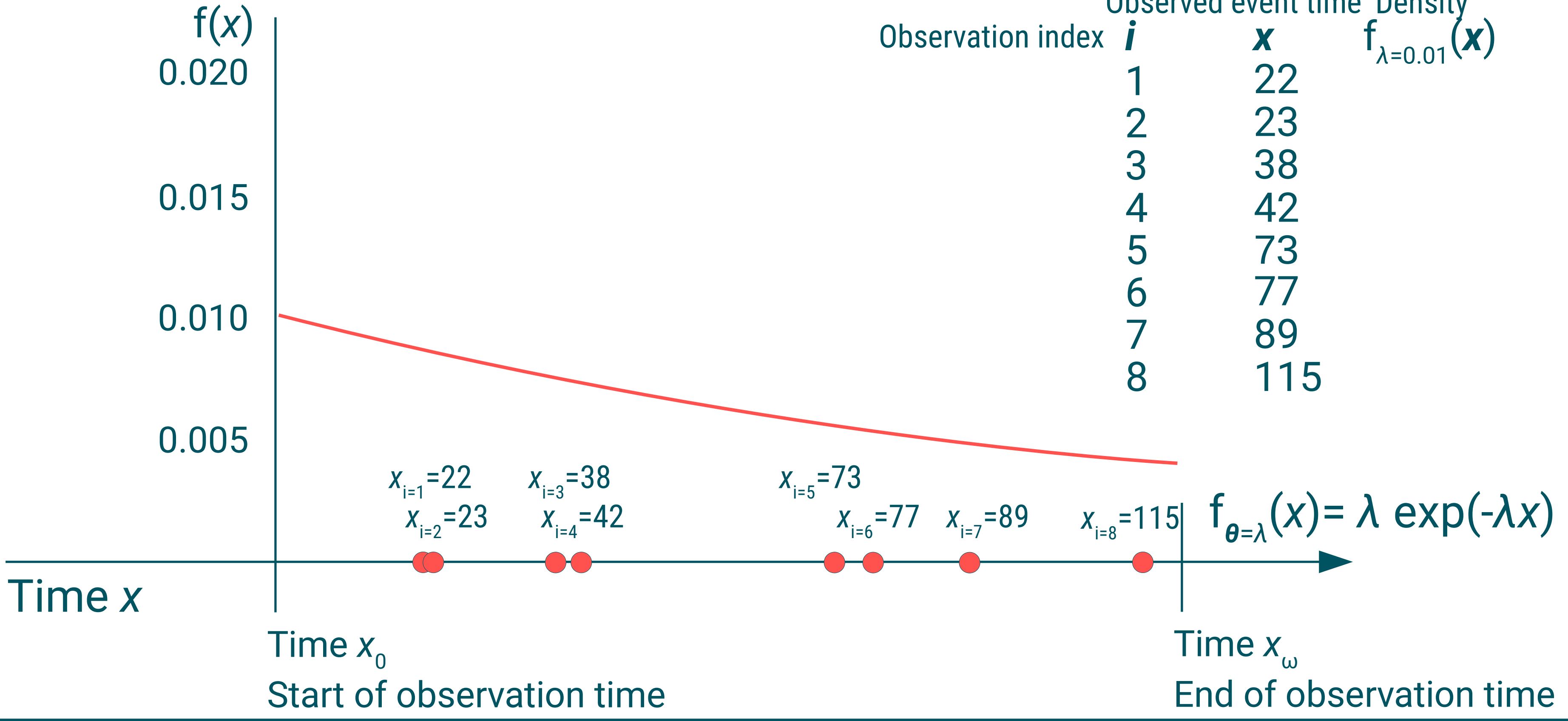
$$f_{\theta=\lambda}(x) = \lambda \exp(-\lambda x)$$

Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Observation index i	Observed event time x	Density $f_{\lambda=0.01}(x)$
1	22	
2	23	
3	38	
4	42	
5	73	
6	77	
7	89	
8	115	



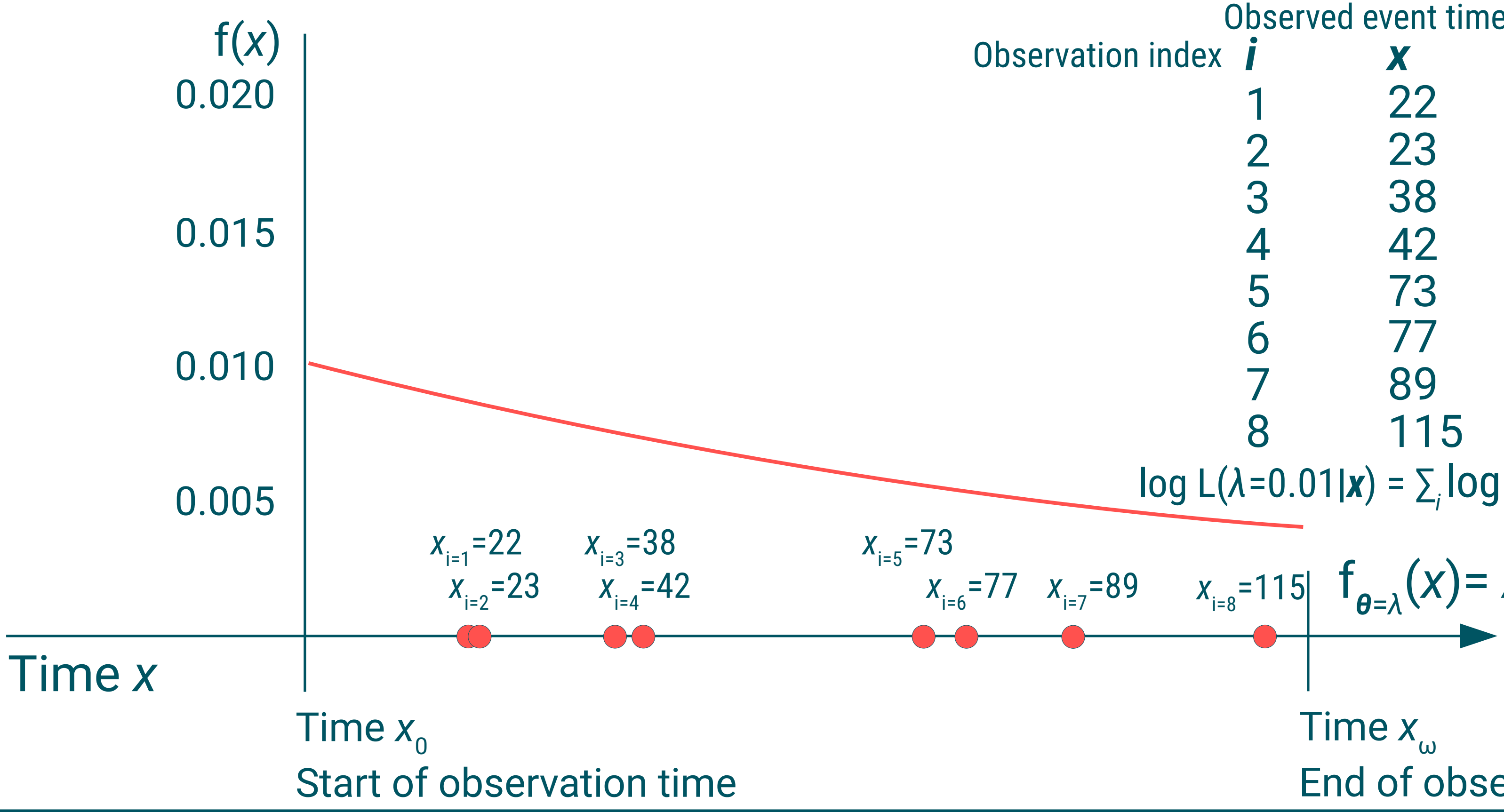
Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Observation index $\boldsymbol{i}$	Observed event time $\boldsymbol{x}$	Density $f_{\lambda=0.01}(\boldsymbol{x})$	log density / log Likelihood contribution $\log f_{\lambda=0.01}(\boldsymbol{x})$
1	22		
2	23		
3	38		
4	42		
5	73		
6	77		
7	89		
8	115		

$\log L(\lambda=0.01|\boldsymbol{x}) = \sum_i \log f_{\lambda=0.01}(x_i) =$ **?**



$f_{\theta=\lambda}(x)= \lambda \exp(-\lambda x)$

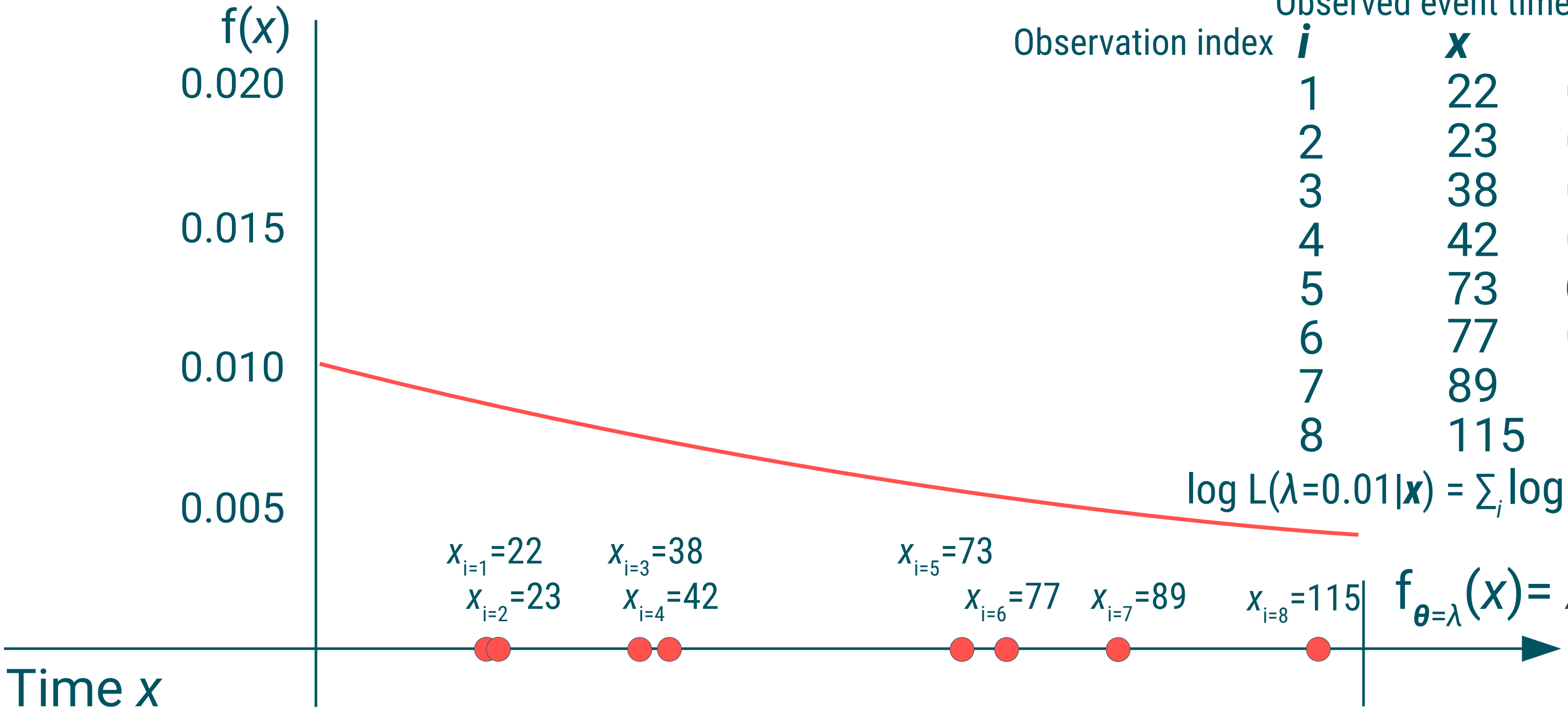
Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Observation index i	Observed event time x	Density $f_{\lambda=0.01}(\boldsymbol{x})$	log density / log Likelihood contribution $\log f_{\lambda=0.01}(\boldsymbol{x})$
1	22	0.0080	-4.83
2	23	0.0079	-4.84
3	38	0.0068	-4.99
4	42	0.0066	-5.03
5	73	0.0048	-5.34
6	77	0.0046	-5.38
7	89	0.0041	-5.50
8	115	0.0032	-5.76

$\log L(\lambda=0.01|\boldsymbol{x}) = \sum_i \log f_{\lambda=0.01}(x_i) =$ **-41.67**



$f_{\boldsymbol{\theta}=\lambda}(x) = \lambda \exp(-\lambda x)$

Time x_0
Start of observation time

Time x_ω
End of observation time

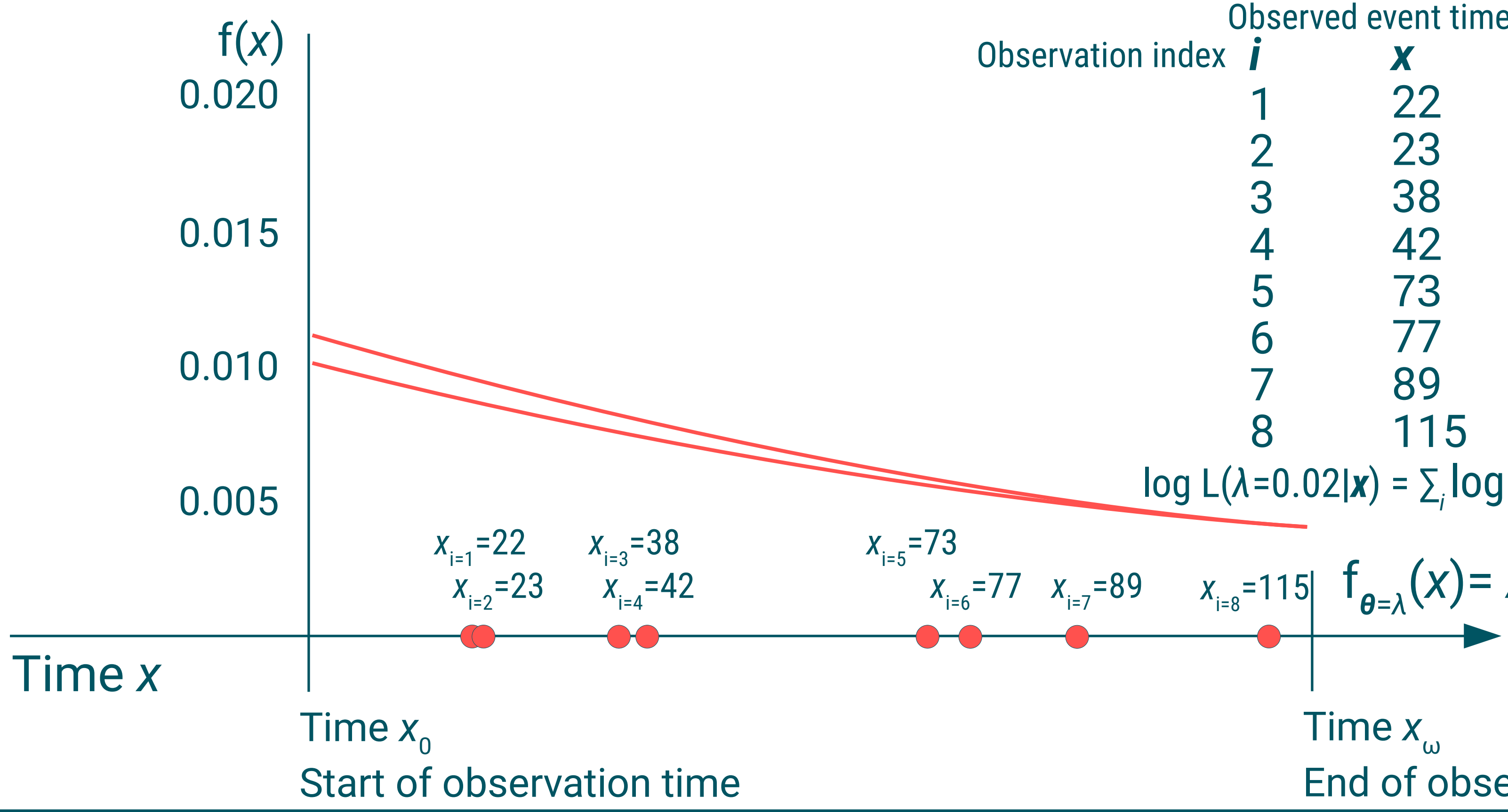
Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Observation index i	Observed event time x	Density $f_{\lambda=0.02}(x)$	log density / log Likelihood contribution $\log f_{\lambda=0.02}(x)$
1	22		
2	23		
3	38		
4	42		
5	73		
6	77		
7	89		
8	115		

$\log L(\lambda=0.02|\boldsymbol{x}) = \sum_i \log f_{\lambda=0.01}(x_i) =$ **?**

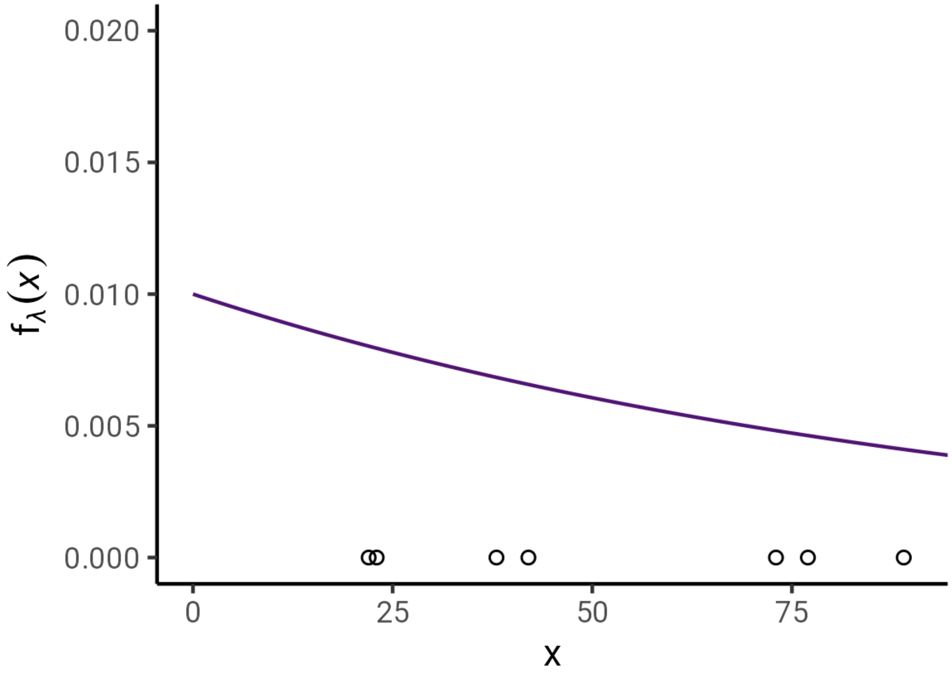


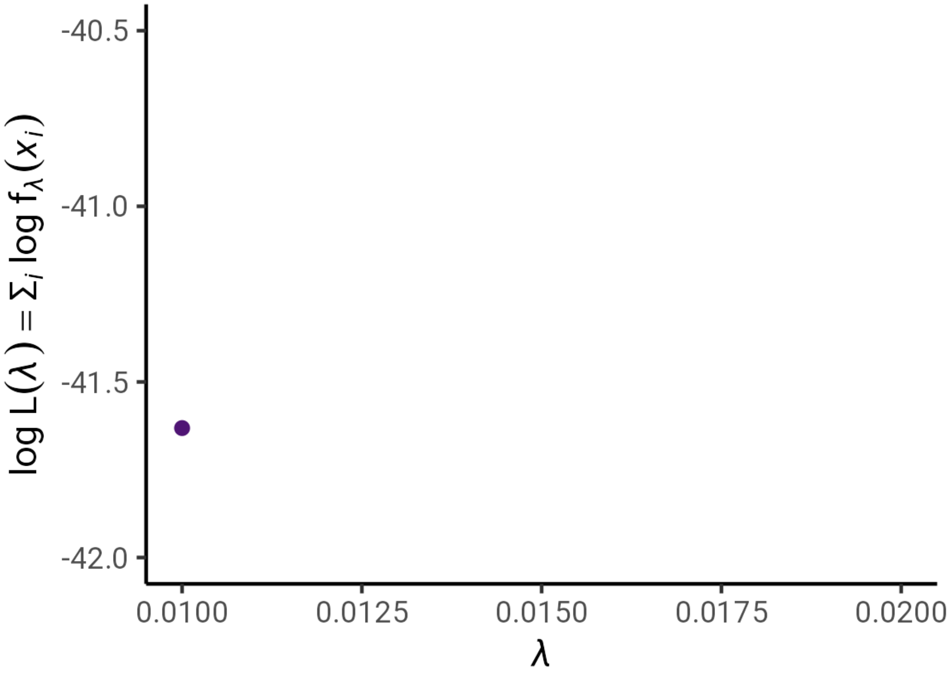
$f_{\theta=\lambda}(x) = \lambda \exp(-\lambda x)$

Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

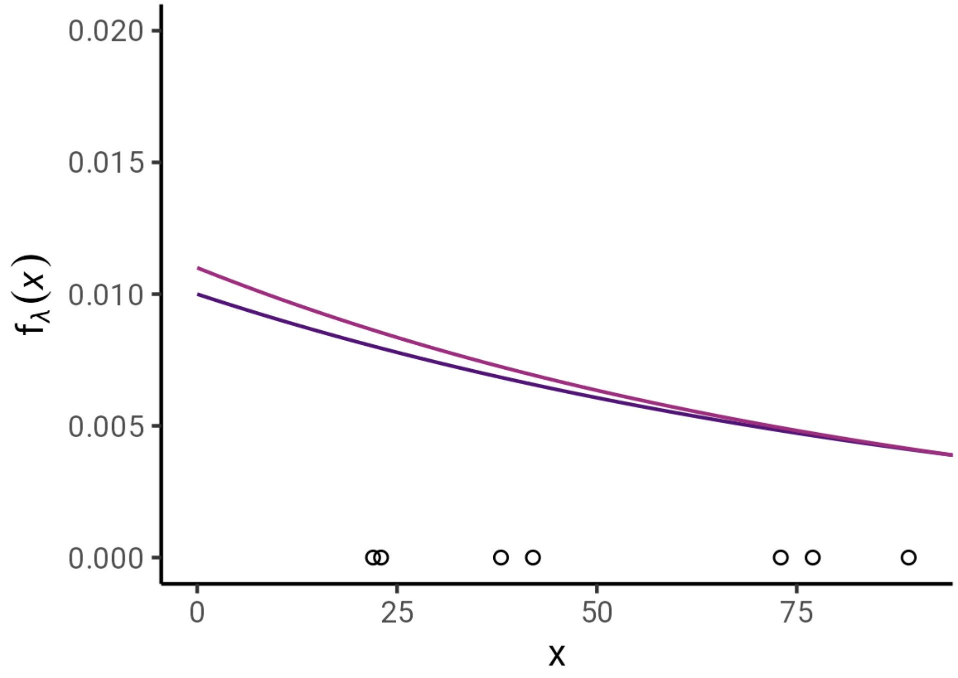


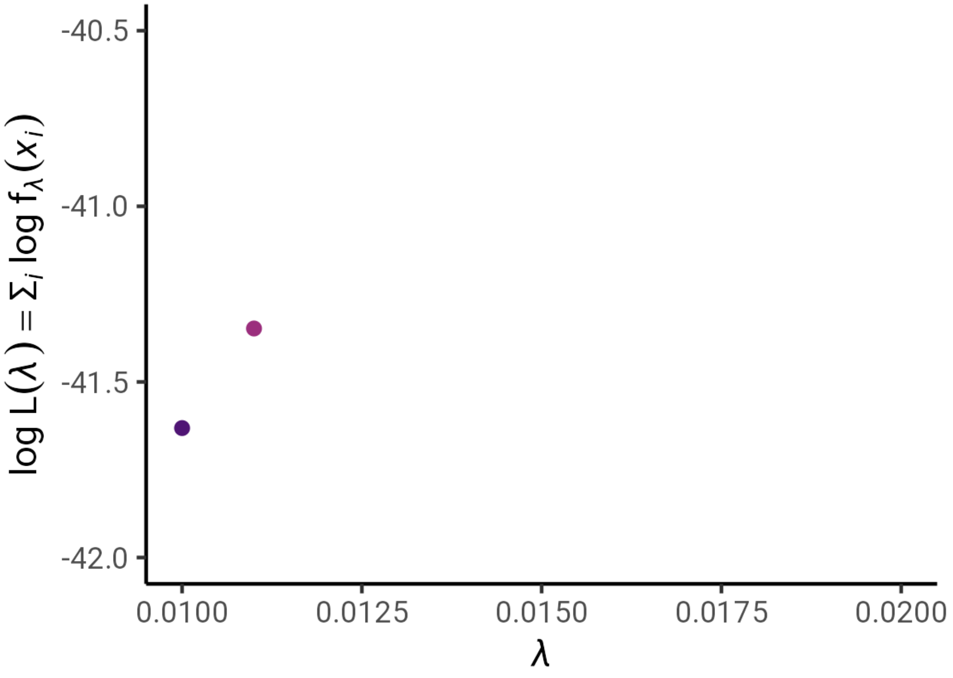


Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

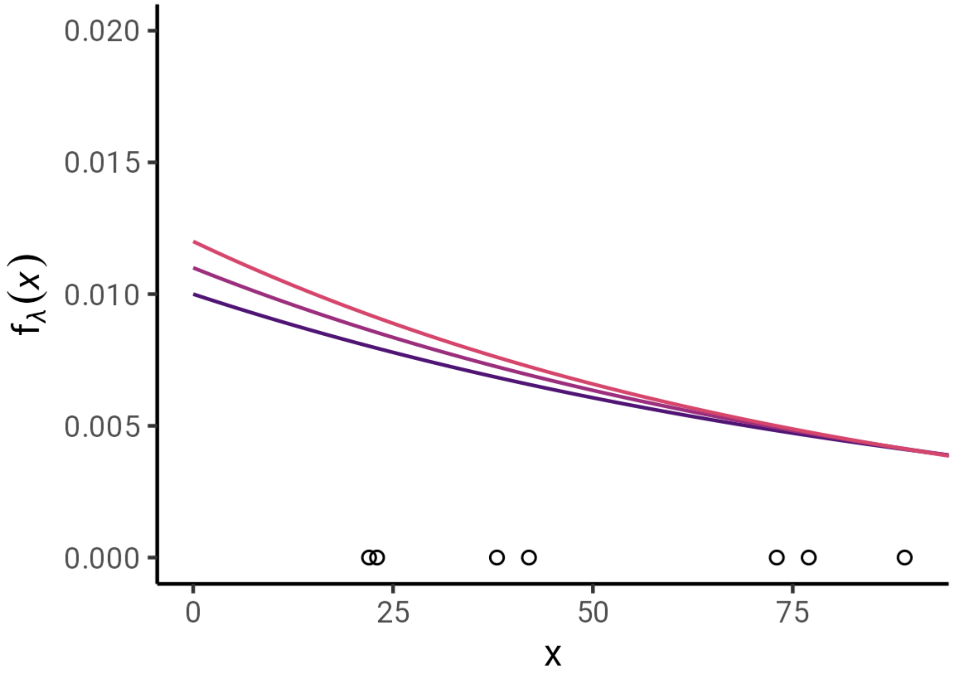


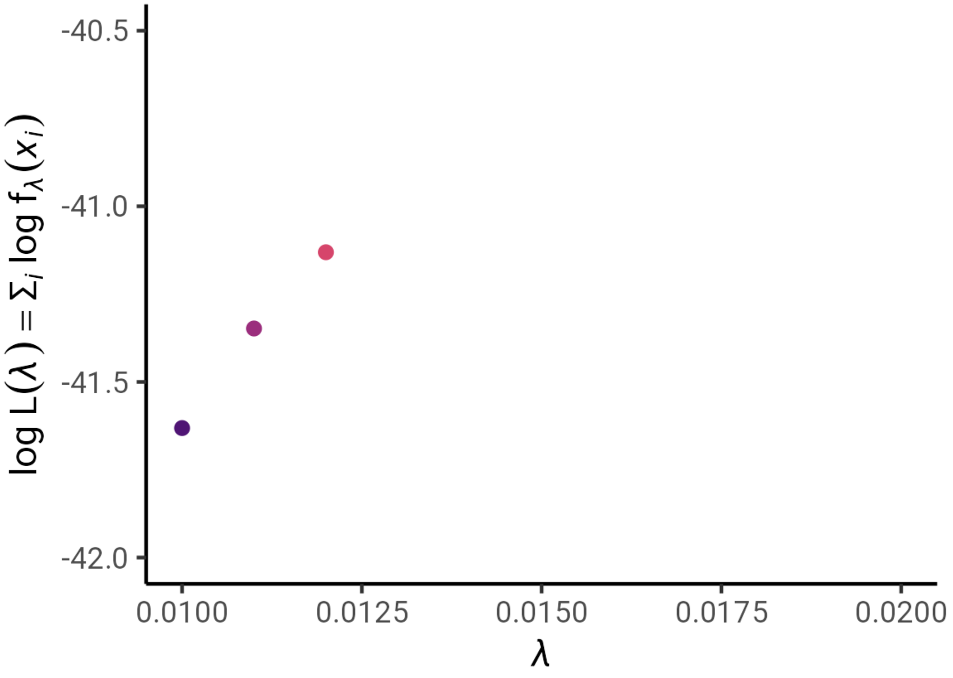


Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

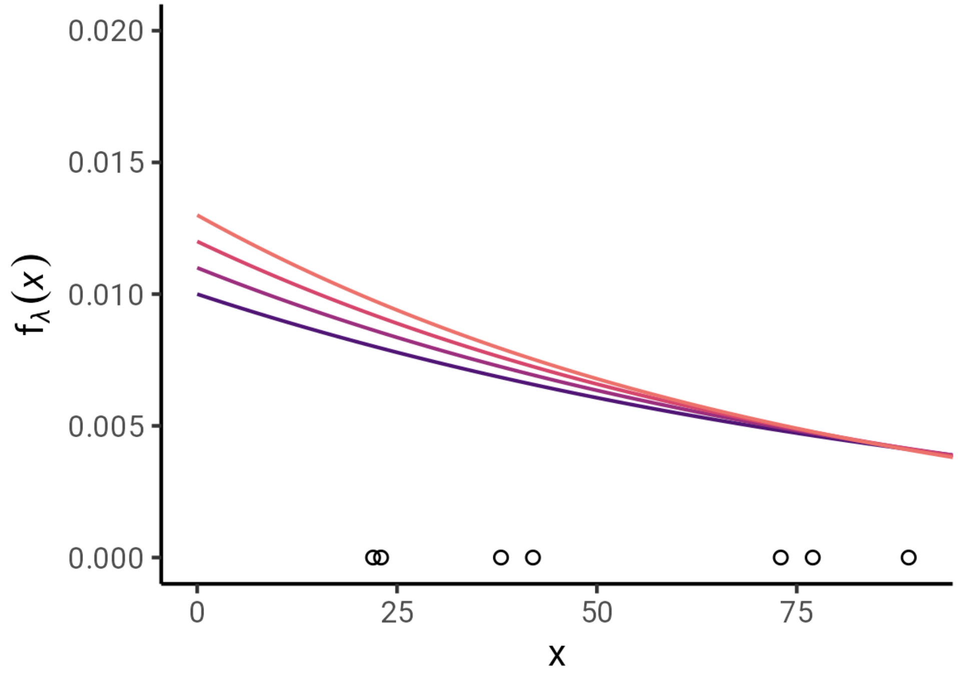
→ Maximum Likelihood Estimation

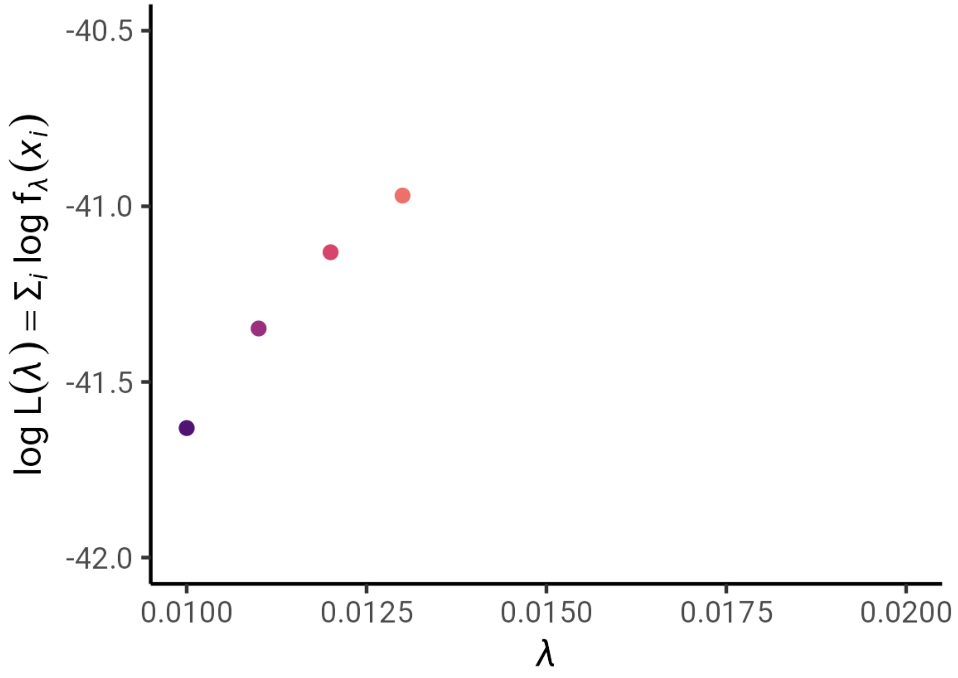




Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?
→ Maximum Likelihood Estimation

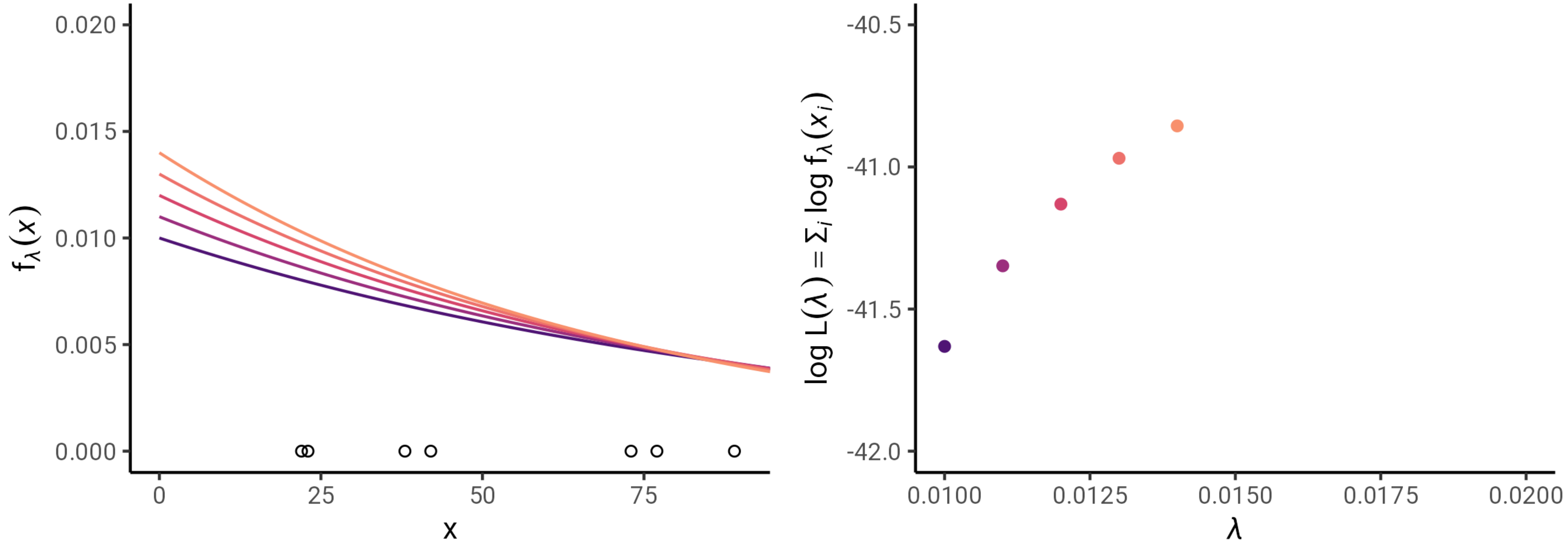




Inferring the Survival Density from Data

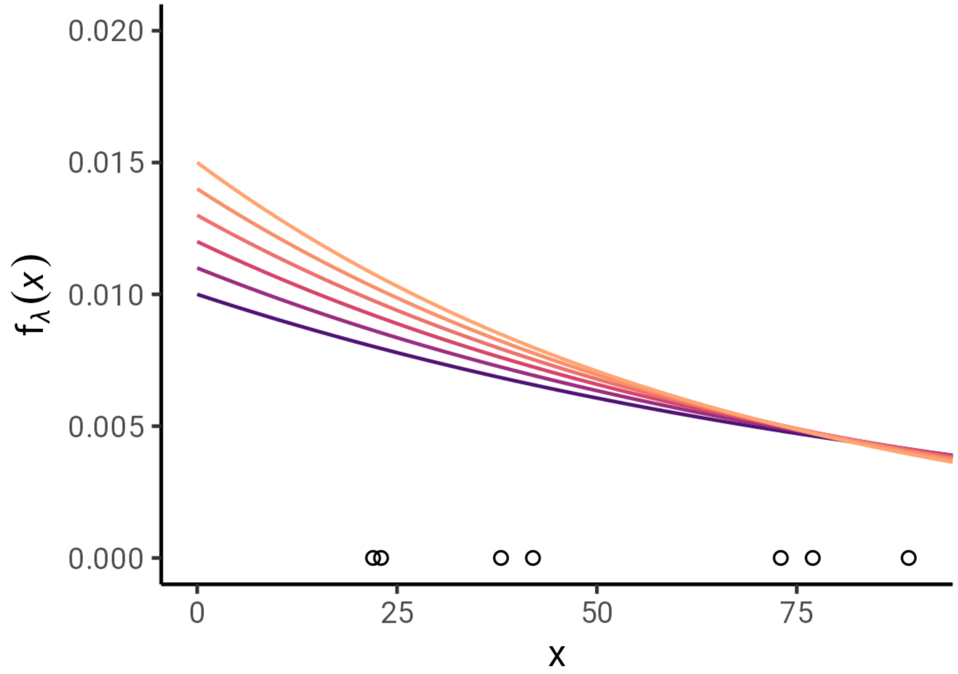
Which distribution most likely corresponds to the data?

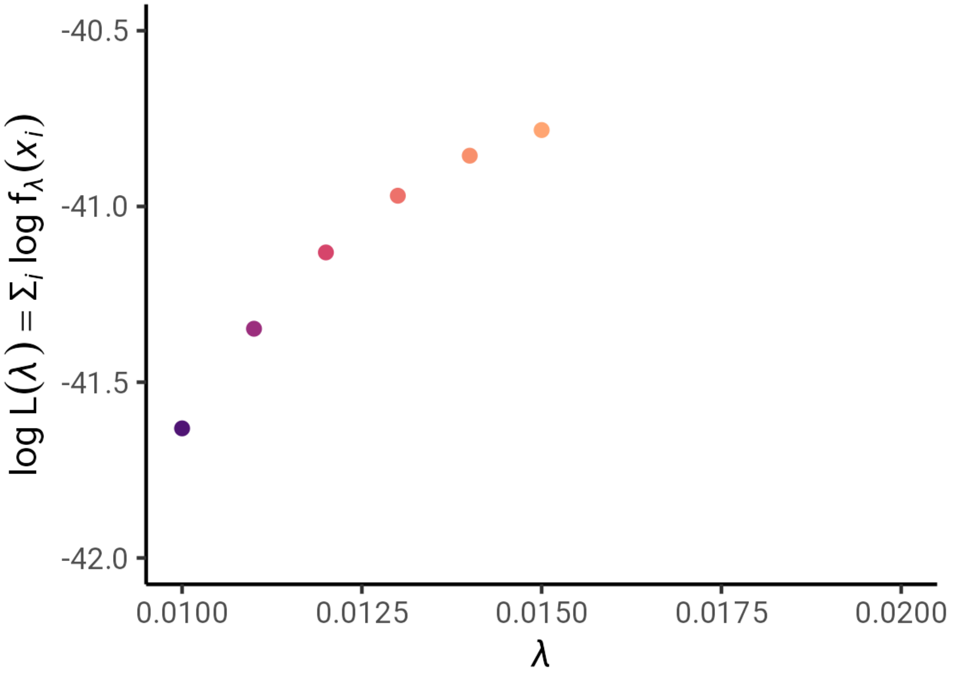
→ Maximum Likelihood Estimation



Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?
→ Maximum Likelihood Estimation

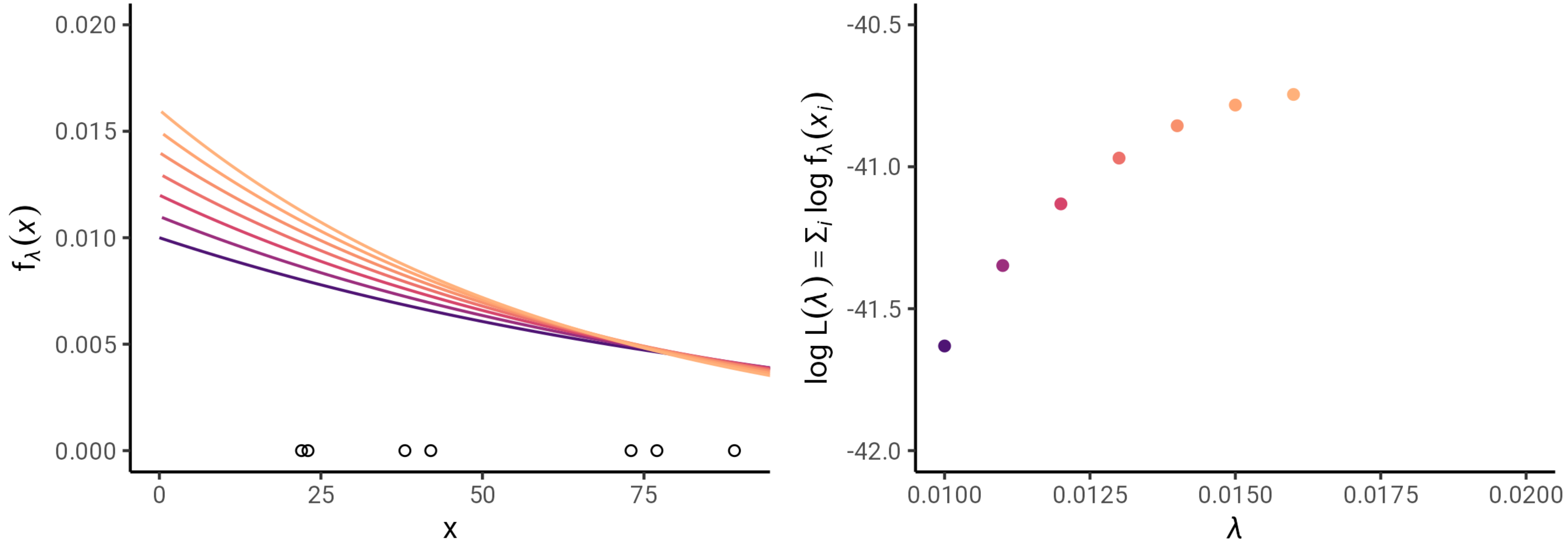




Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

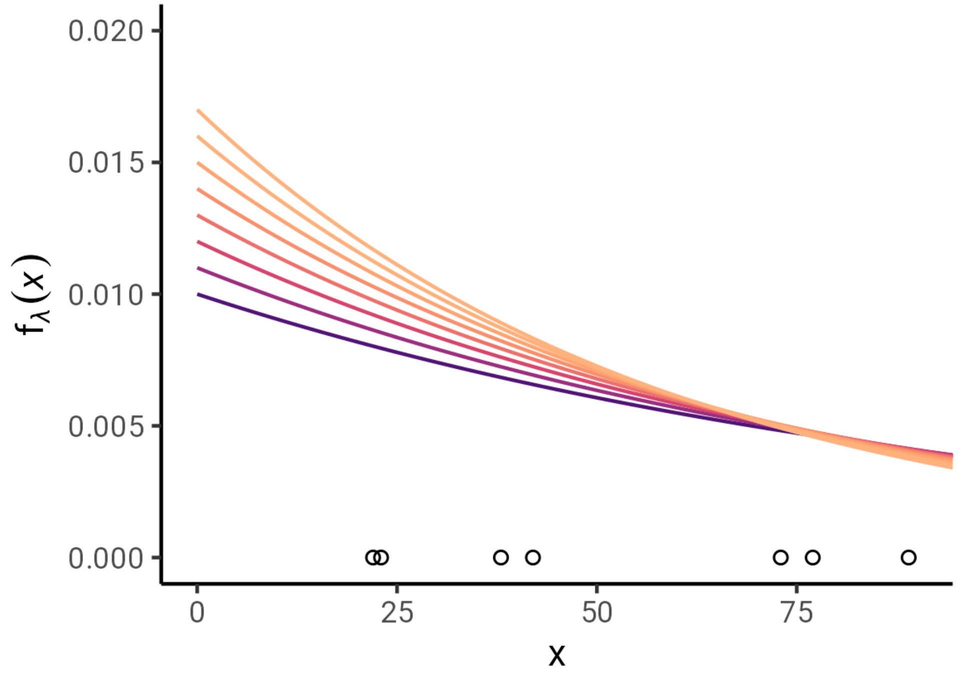
→ Maximum Likelihood Estimation

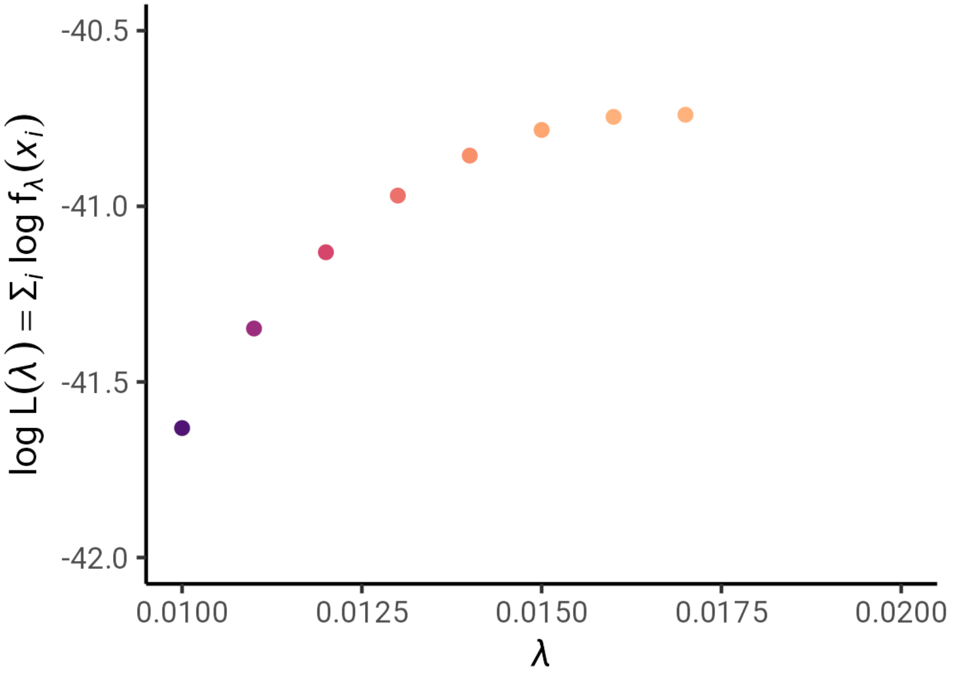


Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

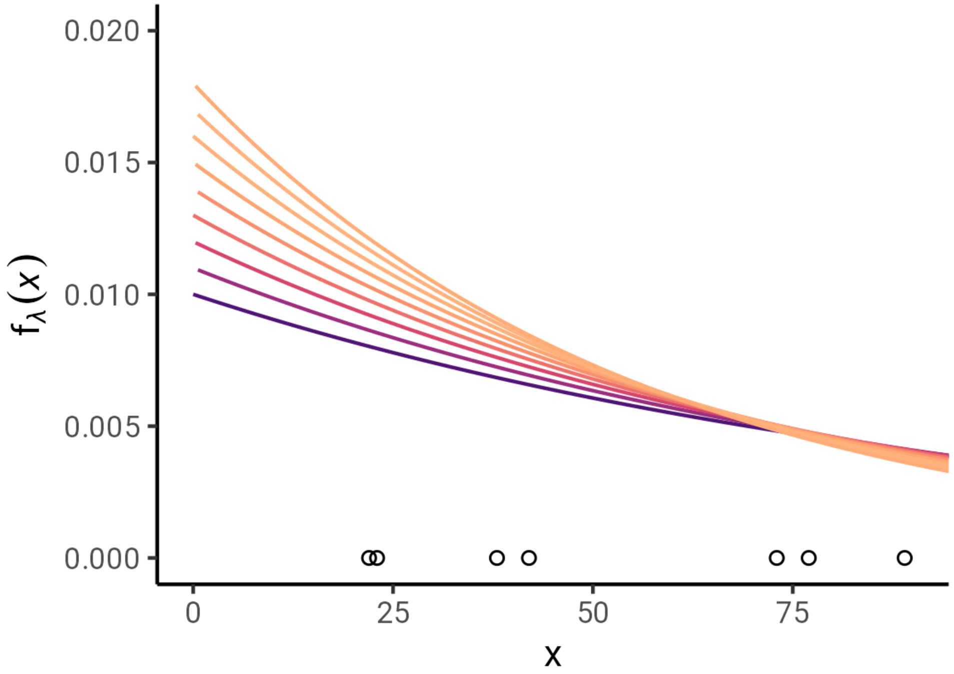
→ Maximum Likelihood Estimation

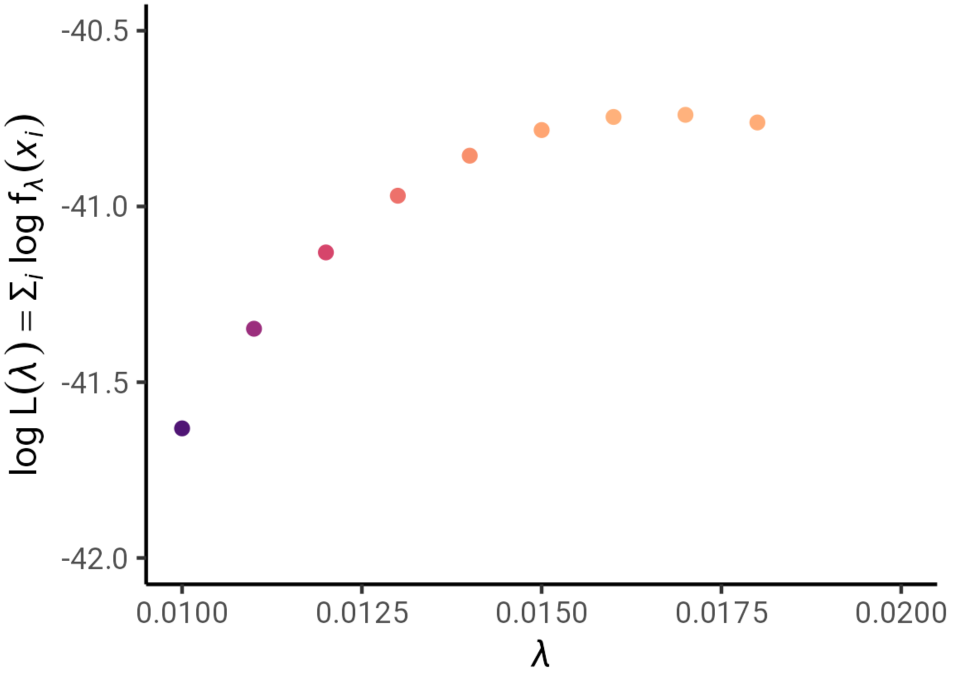




Which distribution most likely corresponds to the data?

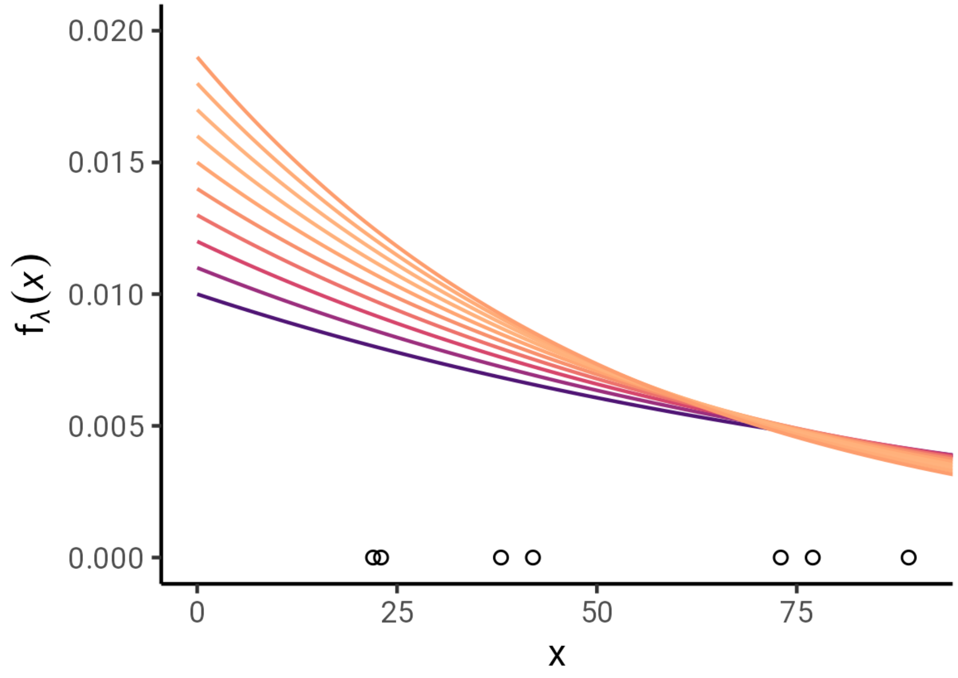
→ Maximum Likelihood Estimation

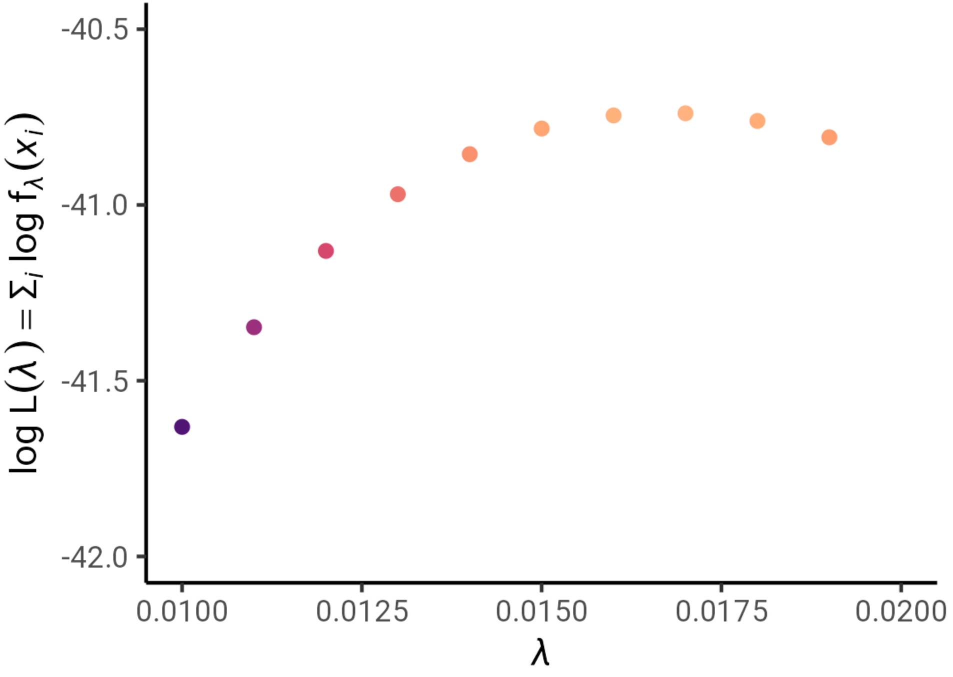




Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?
→ Maximum Likelihood Estimation

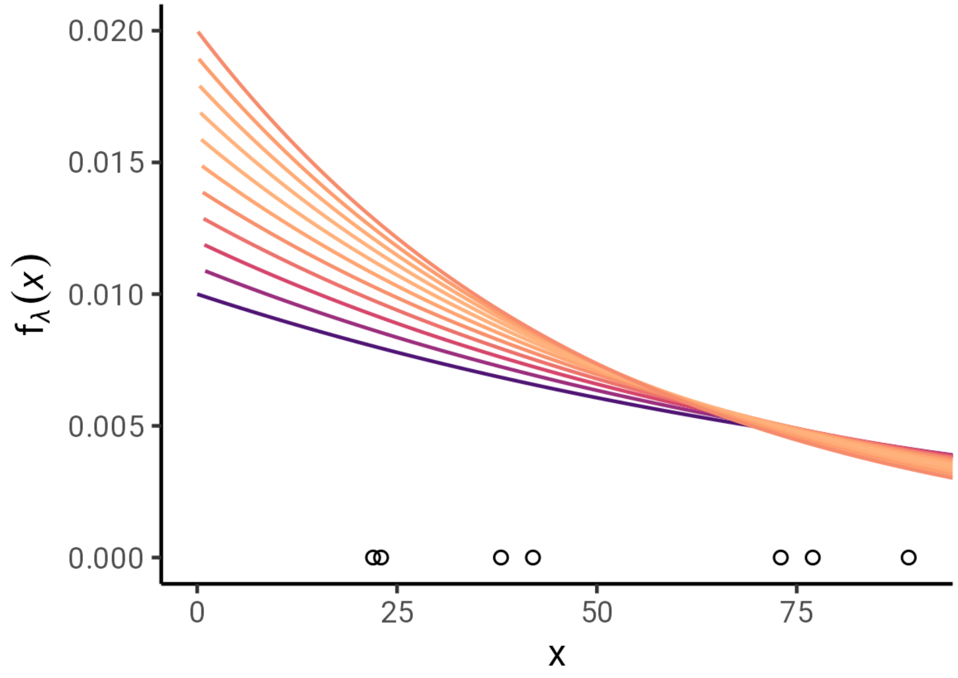


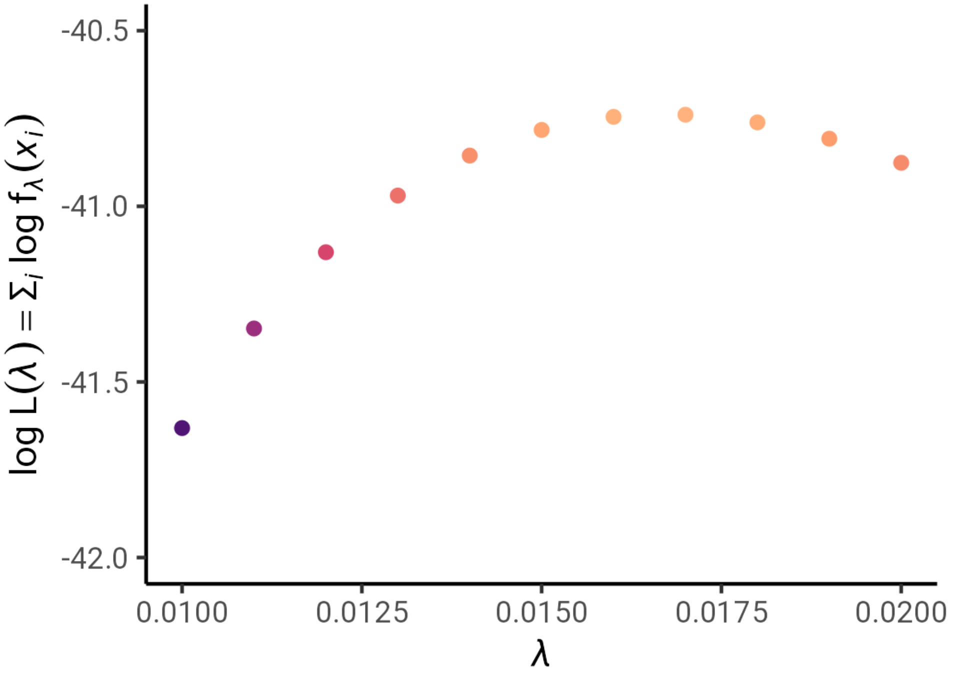


Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation





Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Maximum Likelihood Estimation

We fit a model $f_{\boldsymbol{\theta}}$ to the data $\boldsymbol{x}$ by choosing model parameters $\boldsymbol{\theta}$ which maximize the **likelihood function** L, i.e. which make the observed data most probable.

Product over all observations

Probability density given parameters $\boldsymbol{\theta}$

$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_i f_{\boldsymbol{\theta}}(x_i)$

Single observed survival time

In practice we often maximize the **log-likelihood** for convenience:

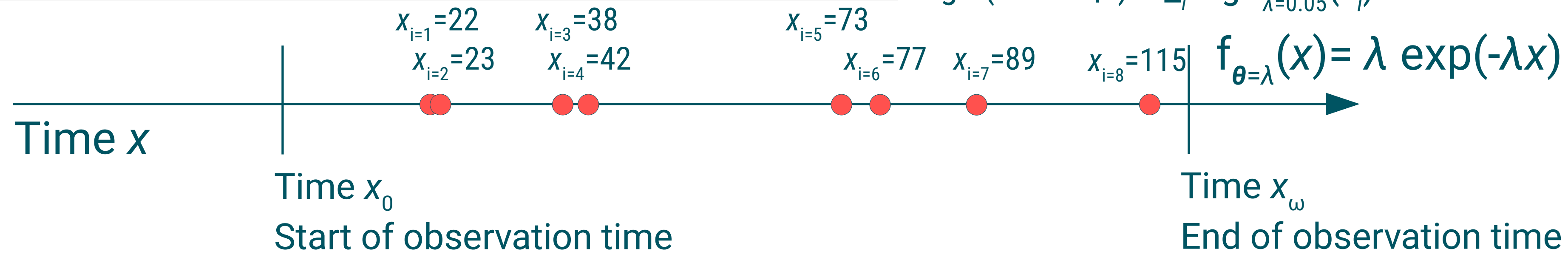
$\log L(\boldsymbol{\theta}|\boldsymbol{x}) = \log \prod_i f_{\boldsymbol{\theta}}(x_i) = \sum_i \log f_{\boldsymbol{\theta}}(x_i)$

Observation index

Observed event time

i	x	$f_{\lambda=0.05}(\boldsymbol{x})$	$\log f_{\lambda=0.05}(\boldsymbol{x})$
1	22	0.0166	-4.095
2	23	0.0158	-4.145
3	38	0.0075	-4.895
4	42	0.0061	-5.095
5	73	0.0013	-6.645
6	77	0.0011	-6.845
7	89	0.0006	-7.744
8	115	0.0002	-8.745

$\log L(\lambda=0.05|\boldsymbol{x}) = \sum_i \log f_{\lambda=0.05}(x_i) = \mathbf{-47.915}$



Censoring

But what if some people did not experience the event during the observation time?

Next week → Censored observations

Materials for this lecture

github.com/jschoeley/survival_analysis-ur-ss22

Jonas Schöley

@jschoeley

0000-0002-3340-8518

j.schoeley@uni-rostock.de