Survival Analysis

Session 3: Incomplete Observations

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…

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*.

h(*x*): Hazard function
The instantaneous rate of new events at time *x* among those who did not experience the event yet.

$$h(x) = \lim_{h\to 0} P(x \le X < x+h \mid X \ge x)/h$$

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

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

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

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

F(*x*): Distribution function
aka Cumulative function
The probability of experiencing the event until time *x*.

$$F(x) = P(X \le x)$$

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

S(*x*): Survival function

The probability of *not* experiencing the event until time *x*.

$$S(x) = P(X > x)$$

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

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

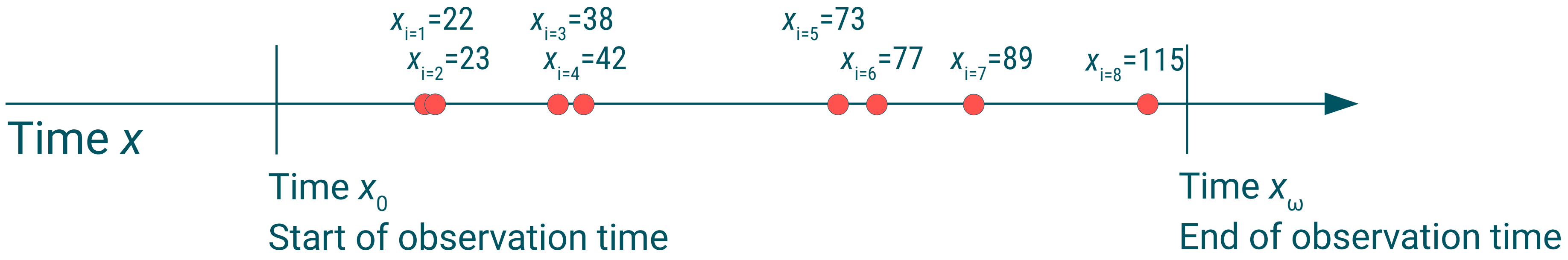
$S(x) = \exp(-H(x))$

What Does Survival Data Look Like?

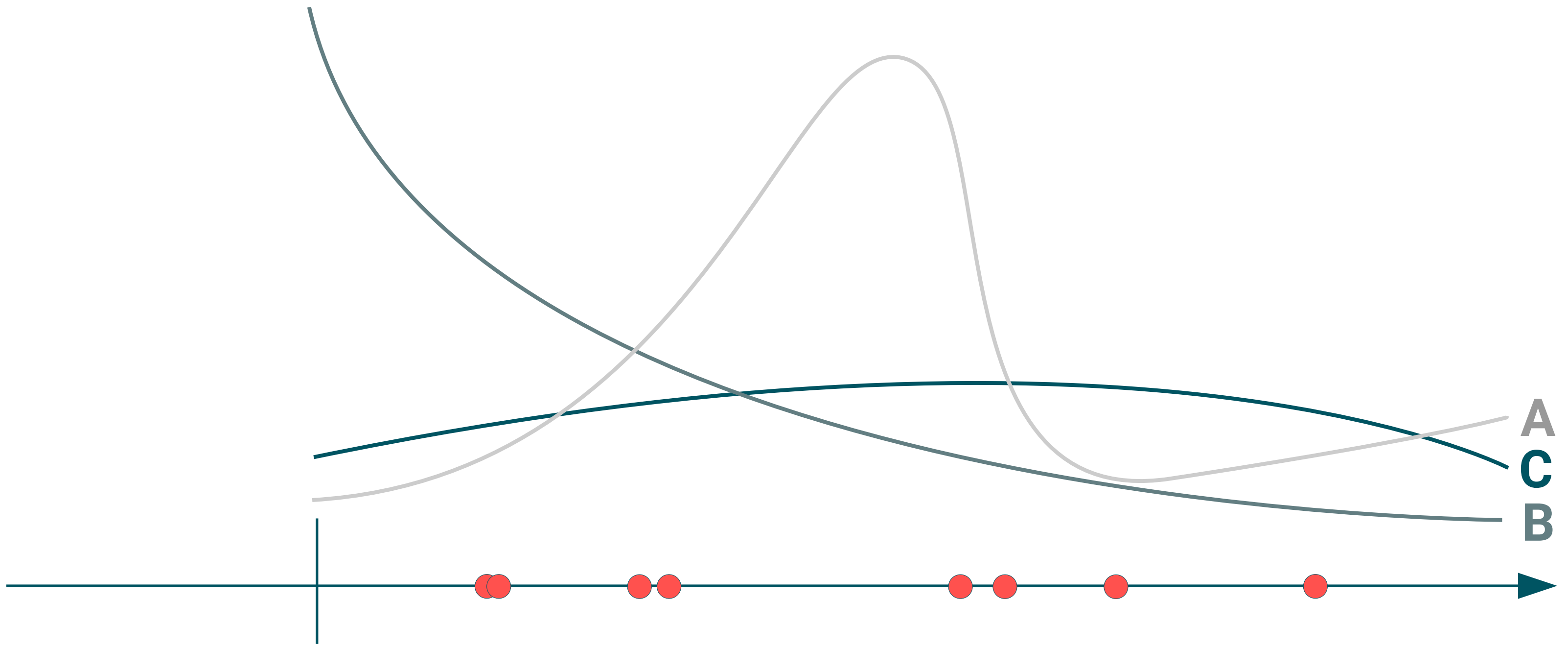
In survival analysis we consider the random variable **“*X*: Time until event”**
x = 0.1 month, 2.3 months…

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

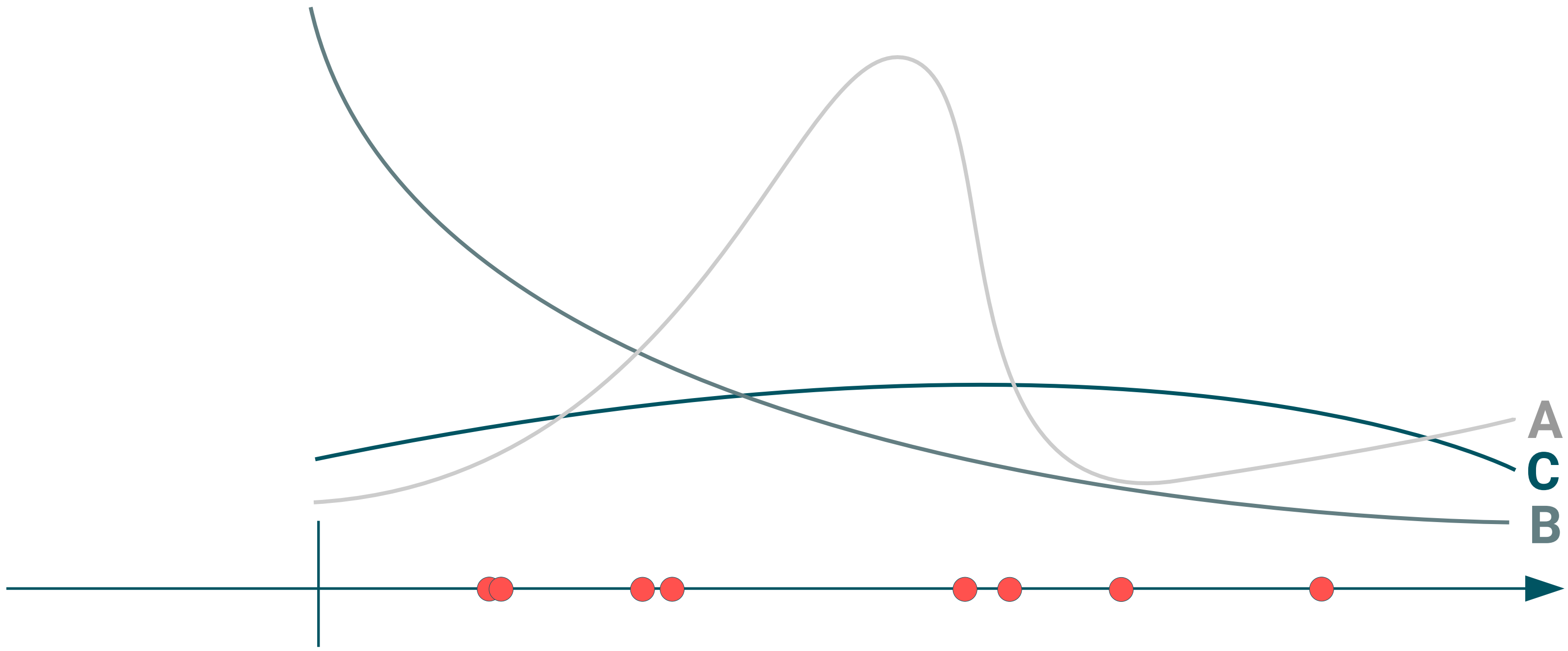


Inferring the Survival Density from Data



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 $\boldsymbol{\theta}$

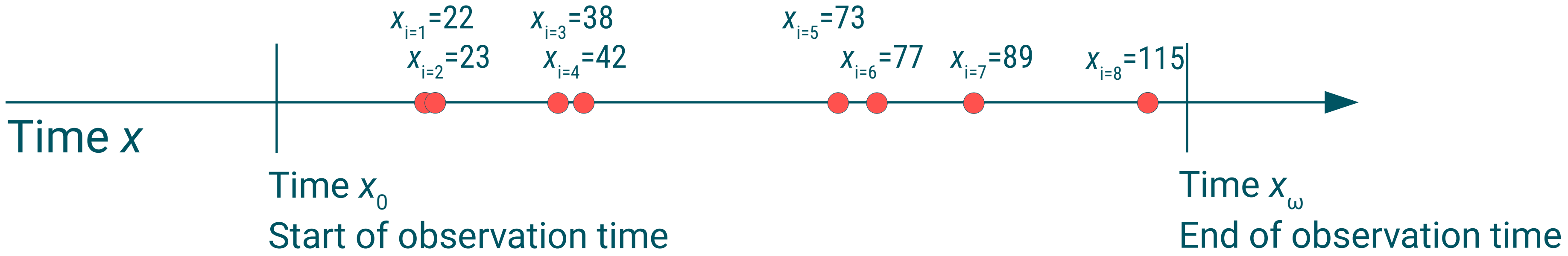
$$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_i f_{\boldsymbol{\theta}}(x_i)$$

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_{θ} 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 $\boldsymbol{\theta}$

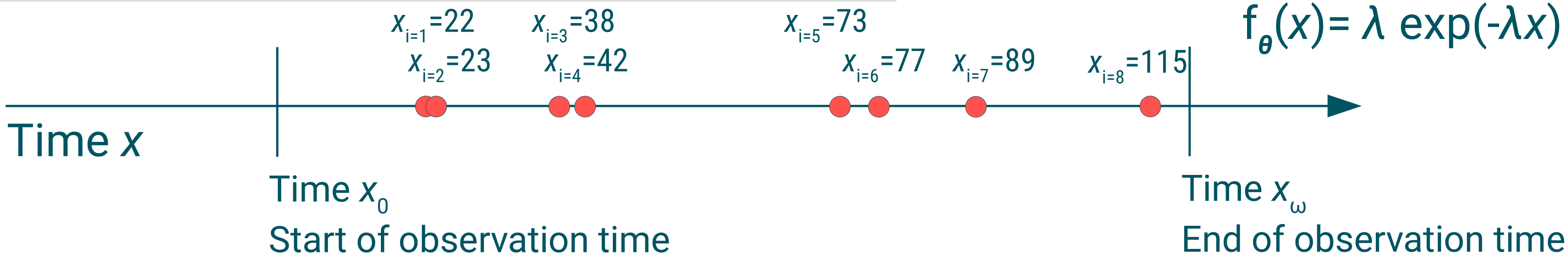
$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_i f_{\boldsymbol{\theta}}(x_i)$

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

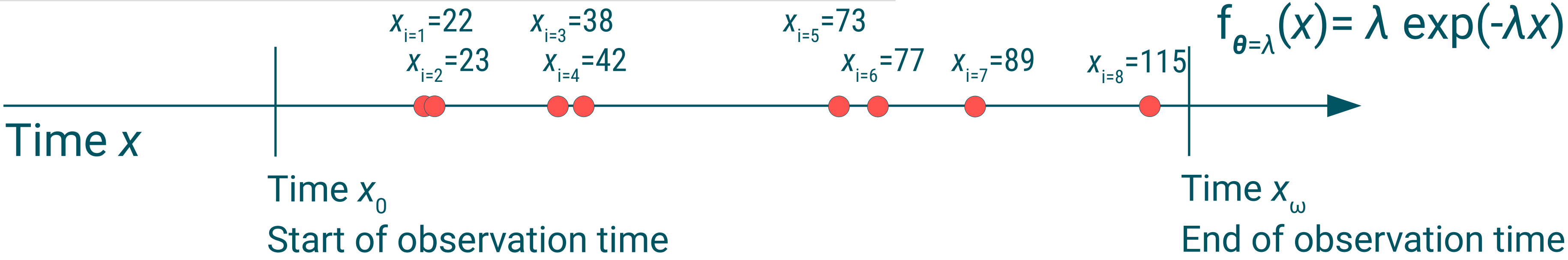
Probability density given parameters $\boldsymbol{\theta}$

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

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

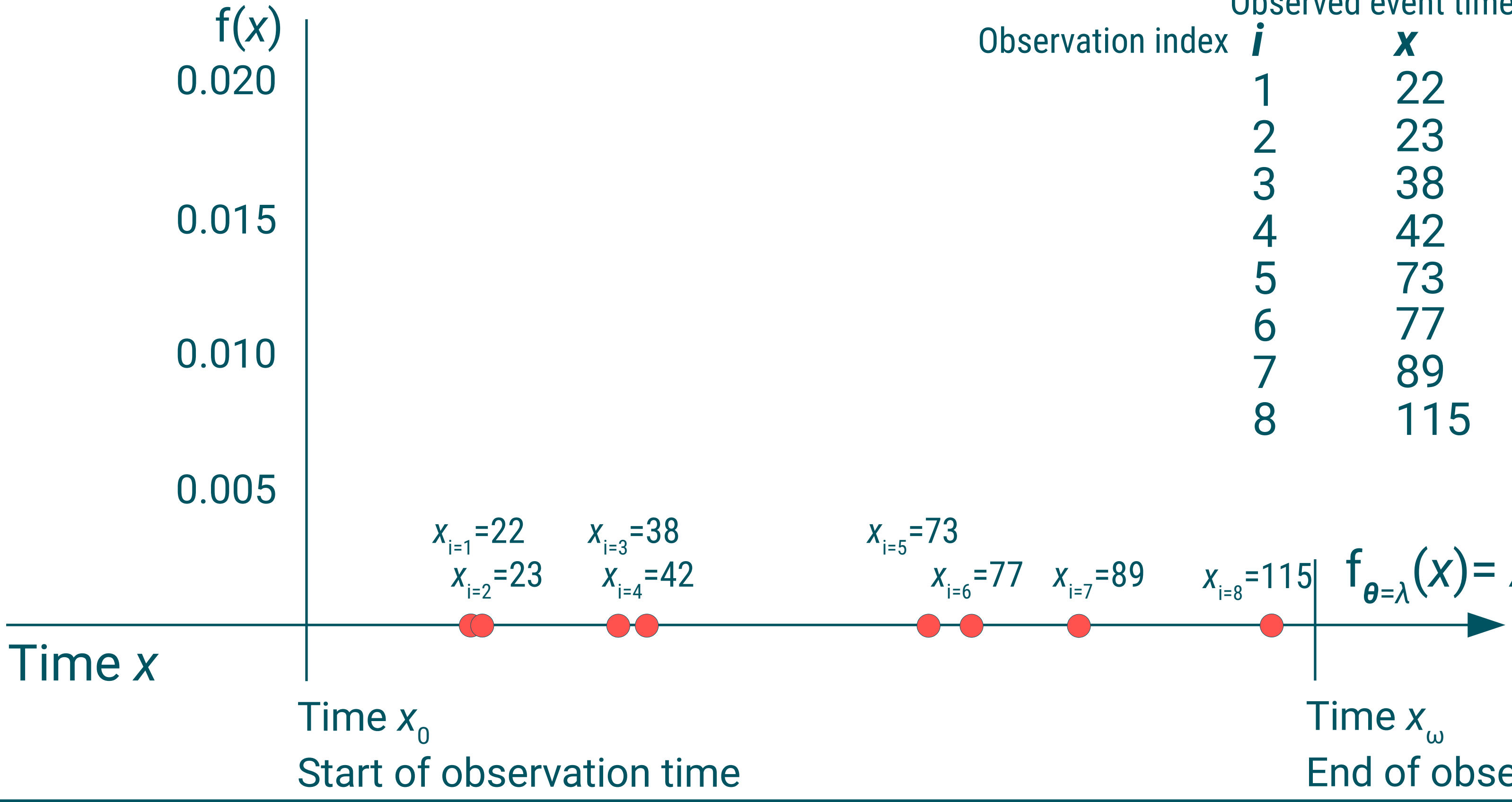


Inferring the Survival Density from Data

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

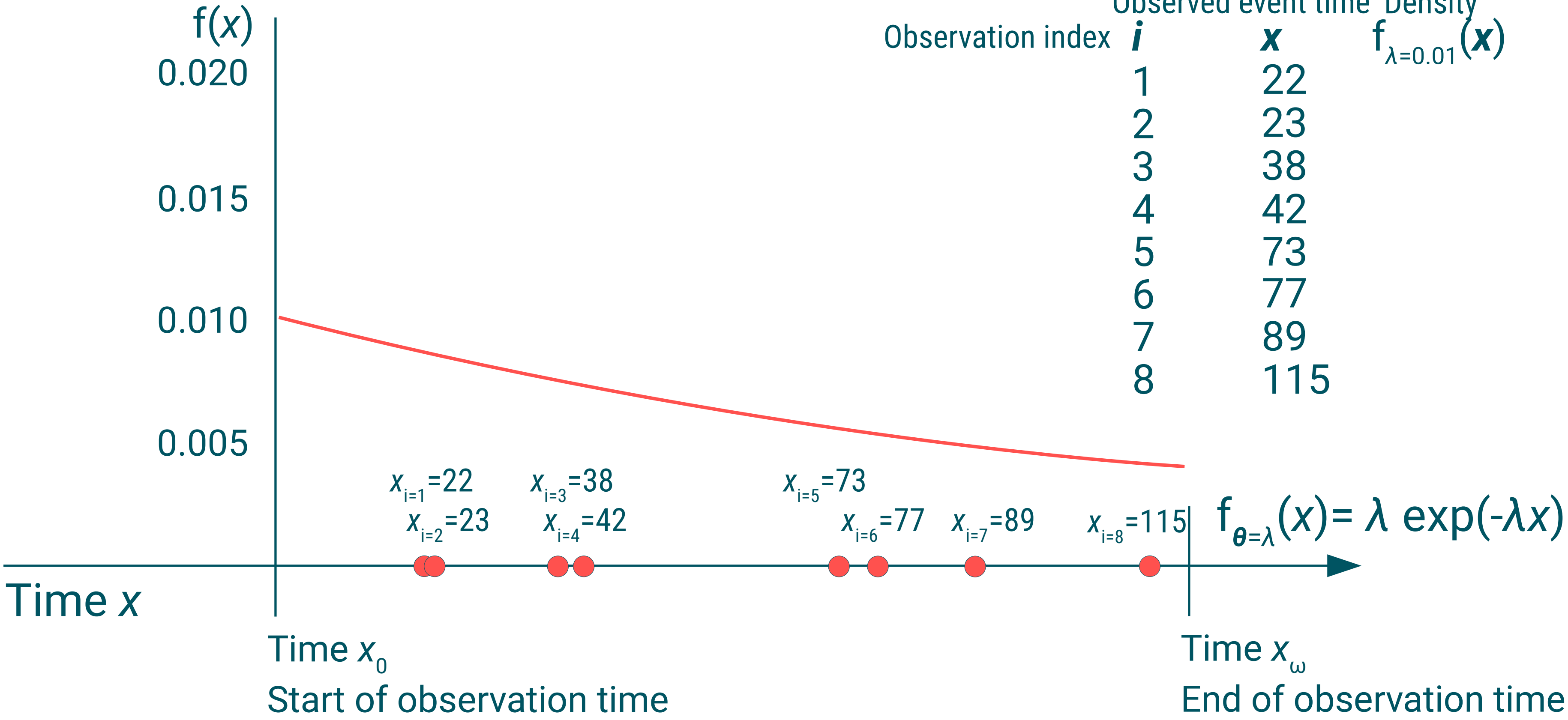


$$f_{\theta=\lambda}(x) = \lambda \exp(-\lambda x)$$

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	

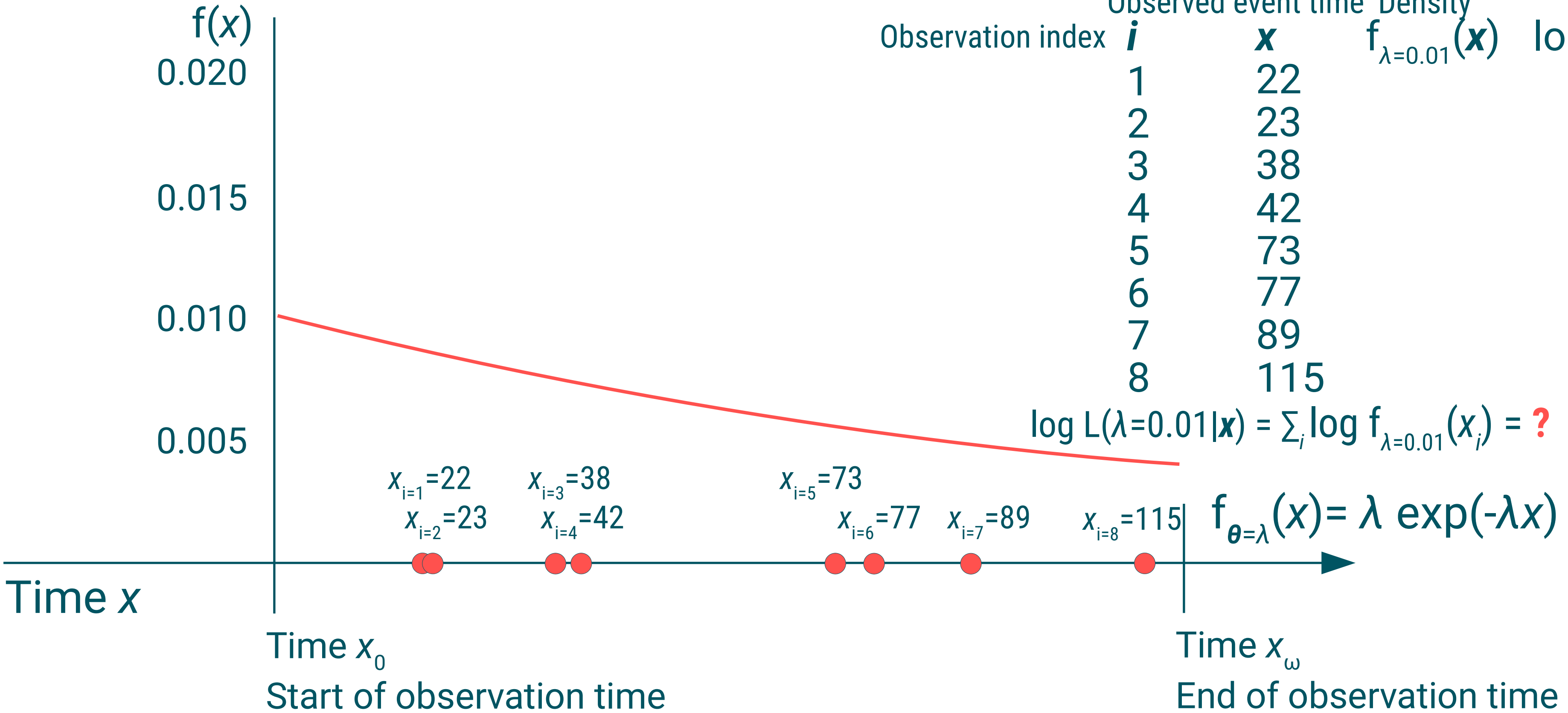
Inferring the Survival Density from Data

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		

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



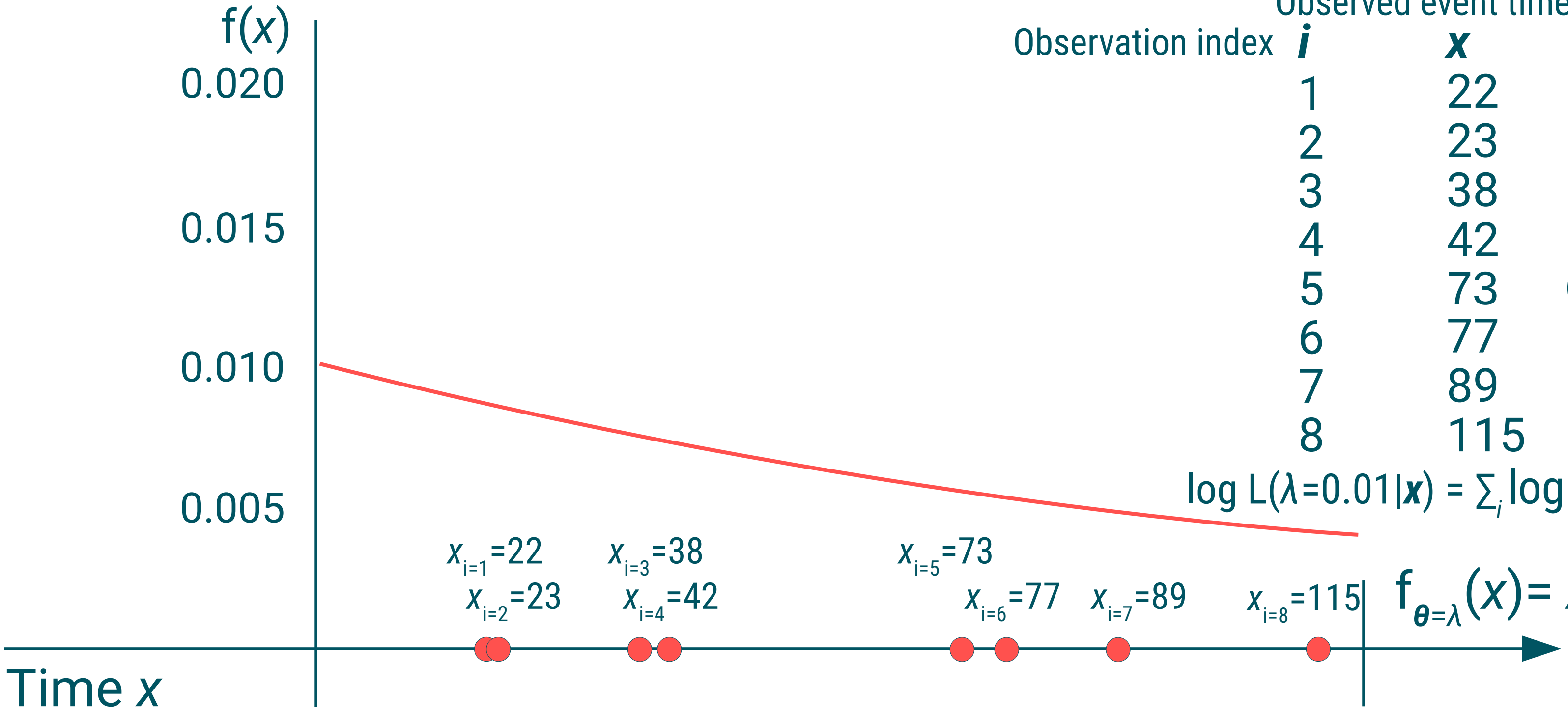
Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

Observation index i	Observed event time x	Density $f_{\lambda=0.01}(x)$	log density / log Likelihood contribution $\log f_{\lambda=0.01}(x)$
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**



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

Time x_0
Start of observation time

Time x_ω
End of observation time

Inferring the Survival Density from Data

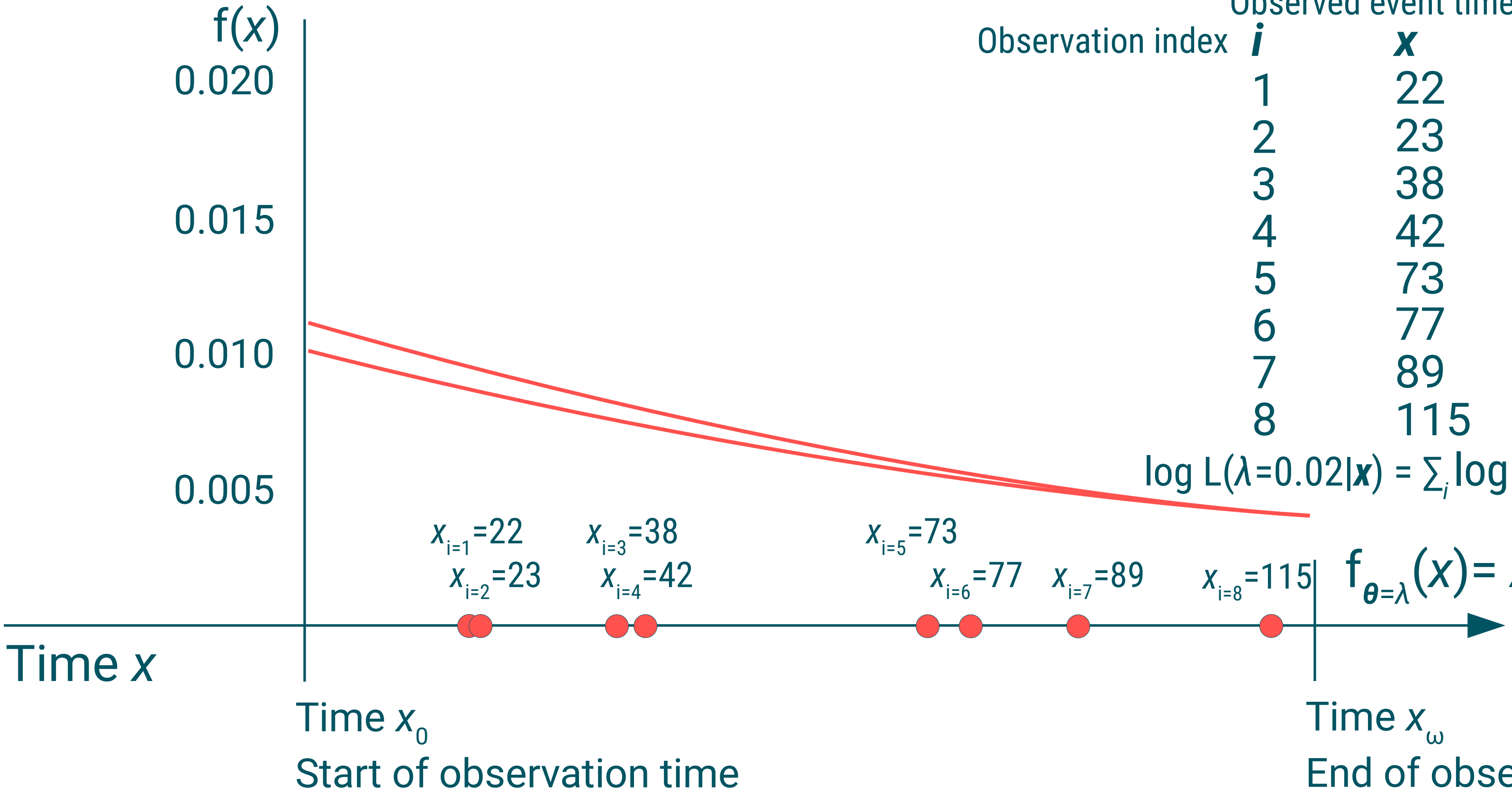
Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

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

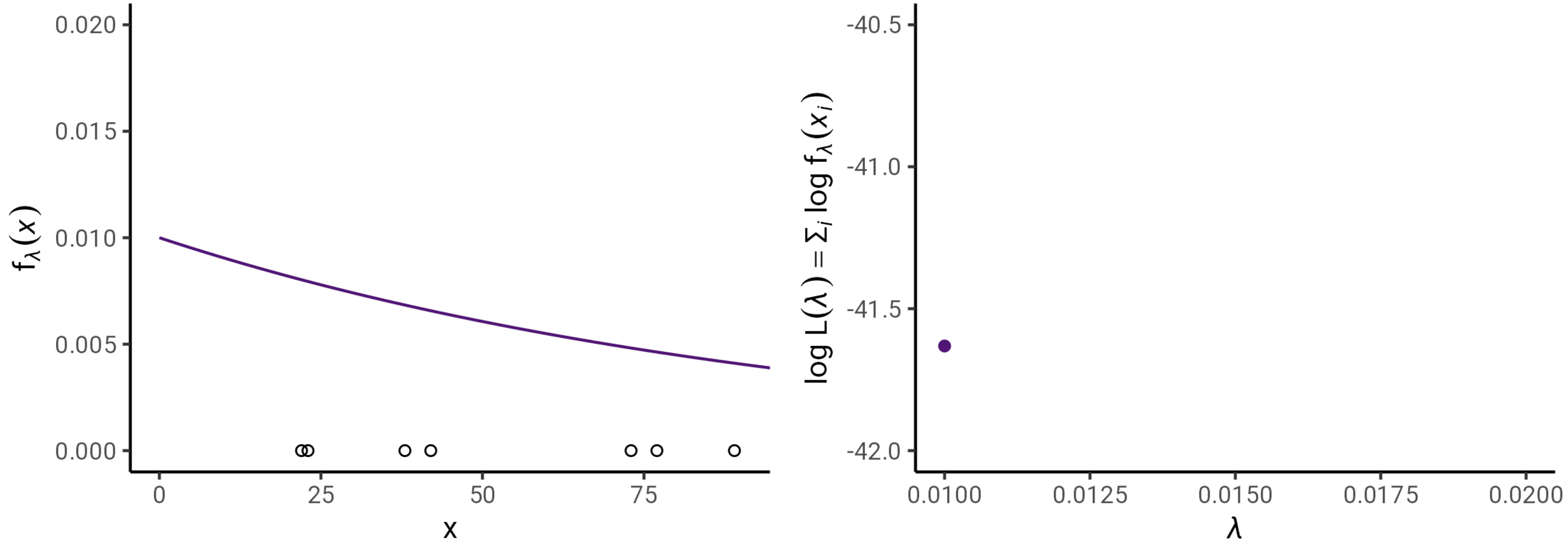
$\log L(\lambda=0.02|\boldsymbol{x}) = \sum_i \log f_{\lambda=0.01}(x_i) = ?$

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



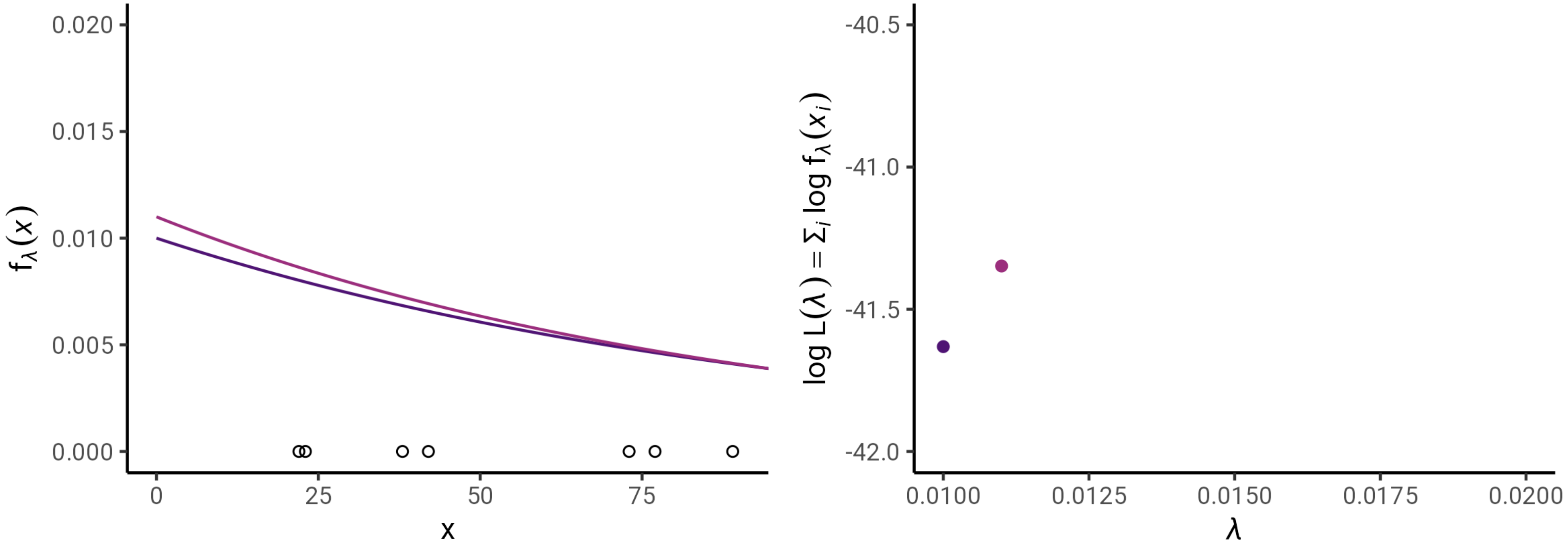
Inferring the Survival Density from Data

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



Which distribution most likely corresponds to the data?

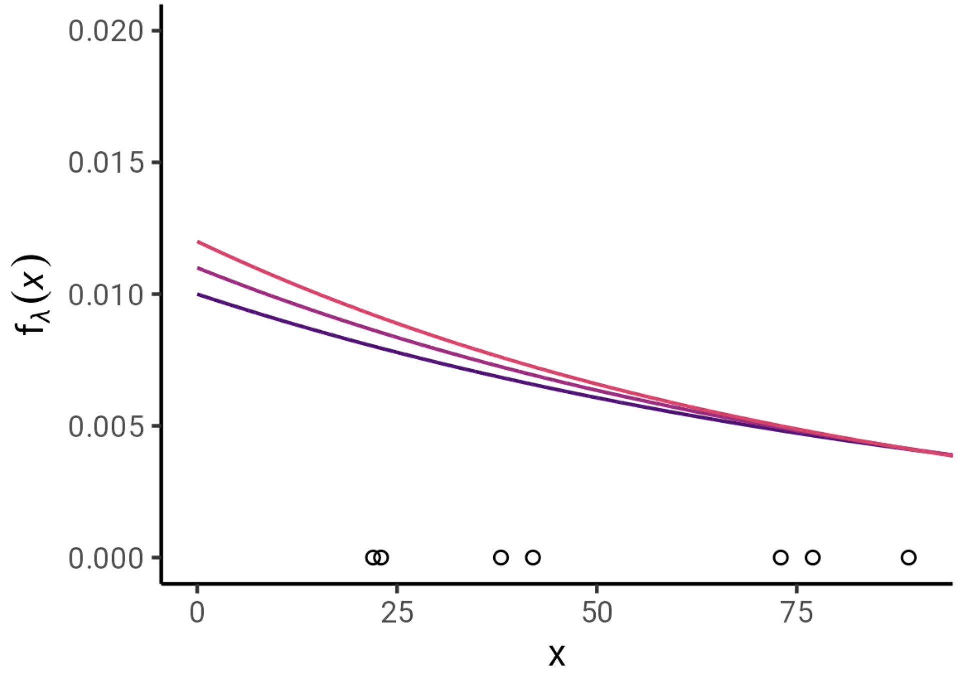
→ Maximum Likelihood Estimation

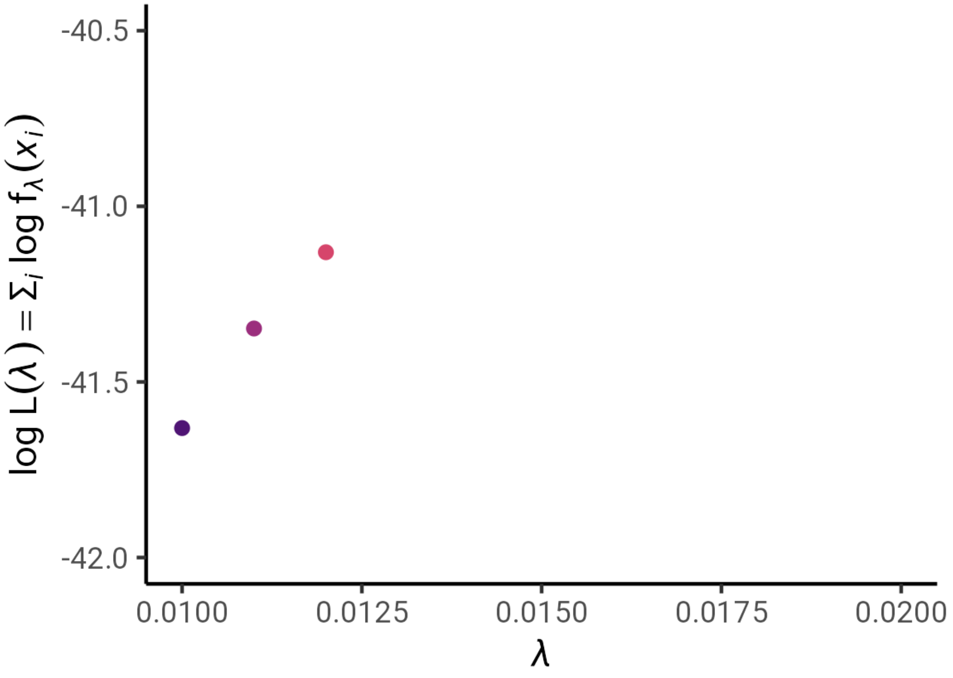


Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

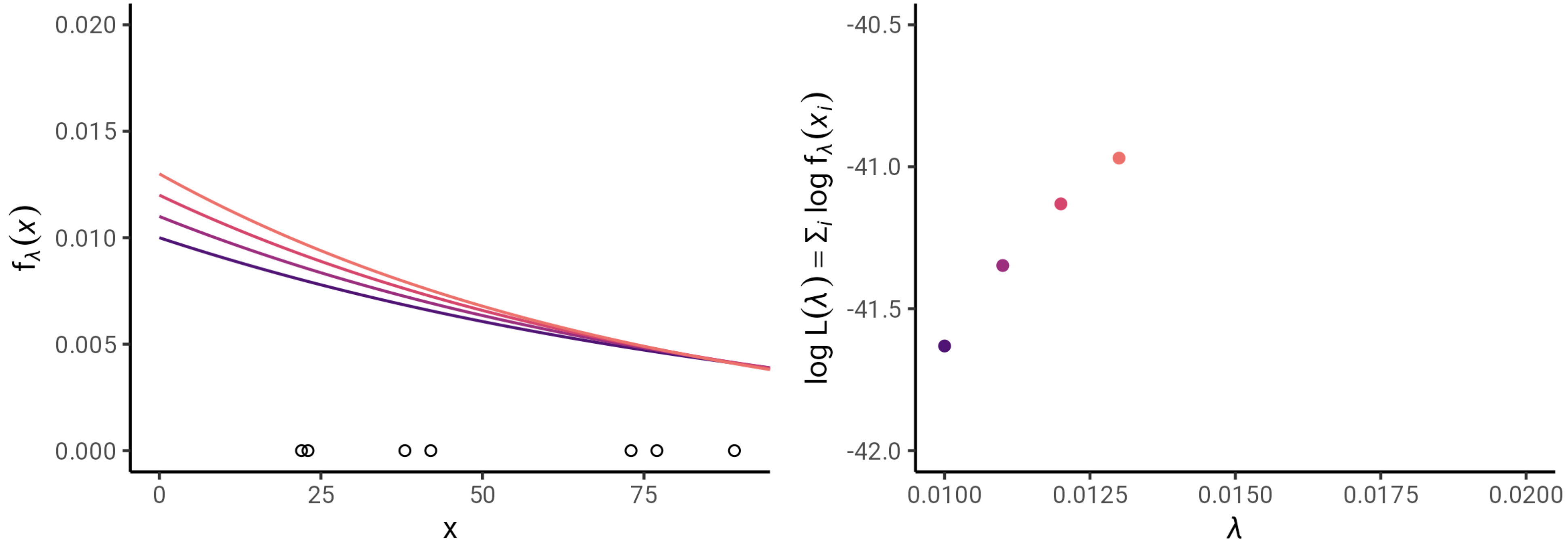
→ Maximum Likelihood Estimation





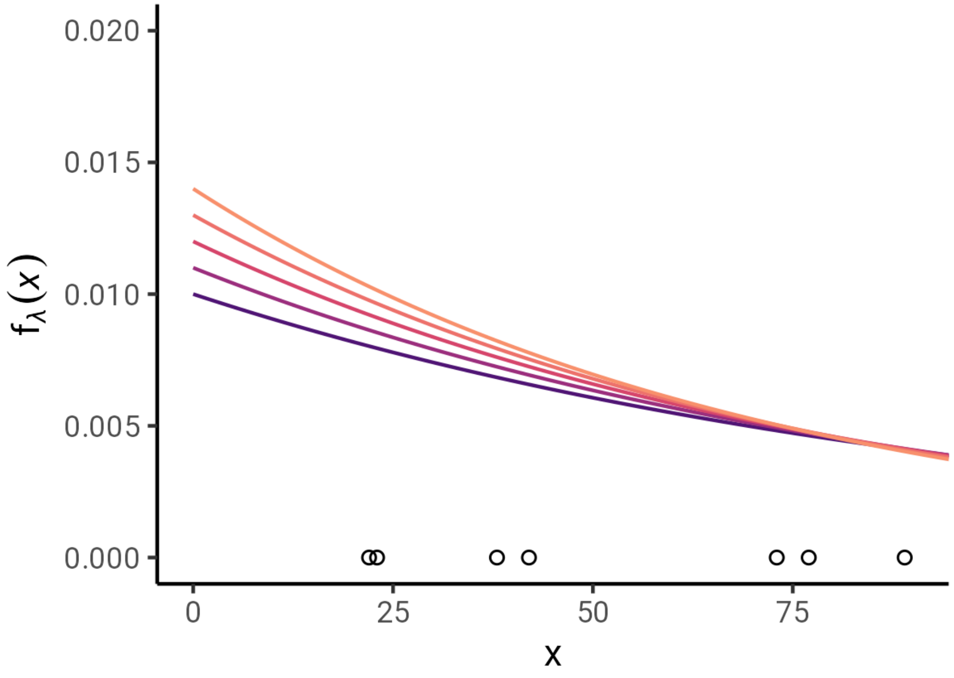
Which distribution most likely corresponds to the data?

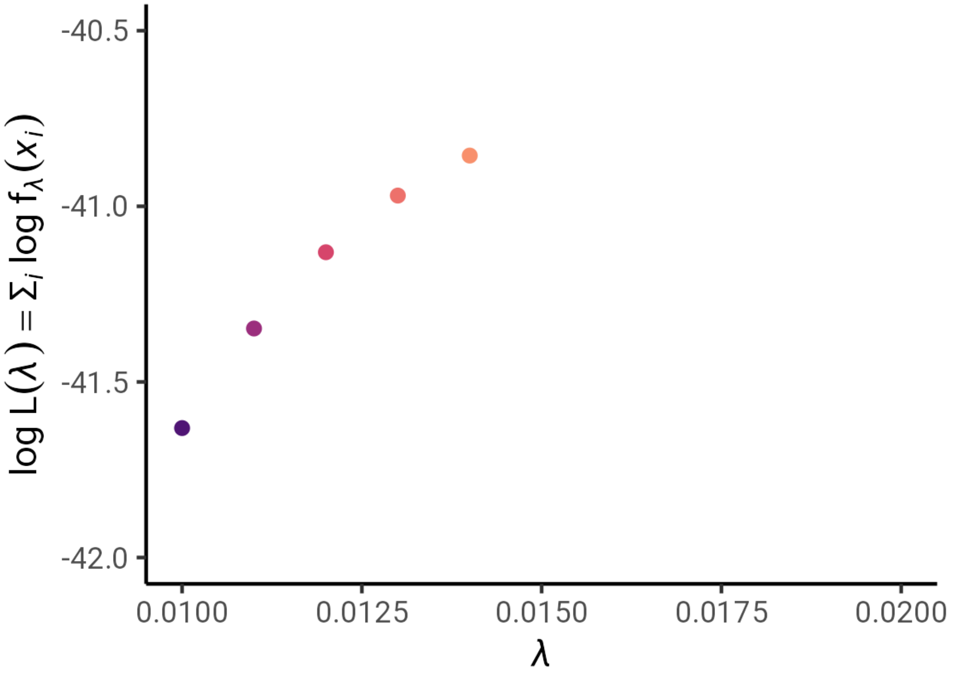
→ Maximum Likelihood Estimation



Inferring the Survival Density from Data

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

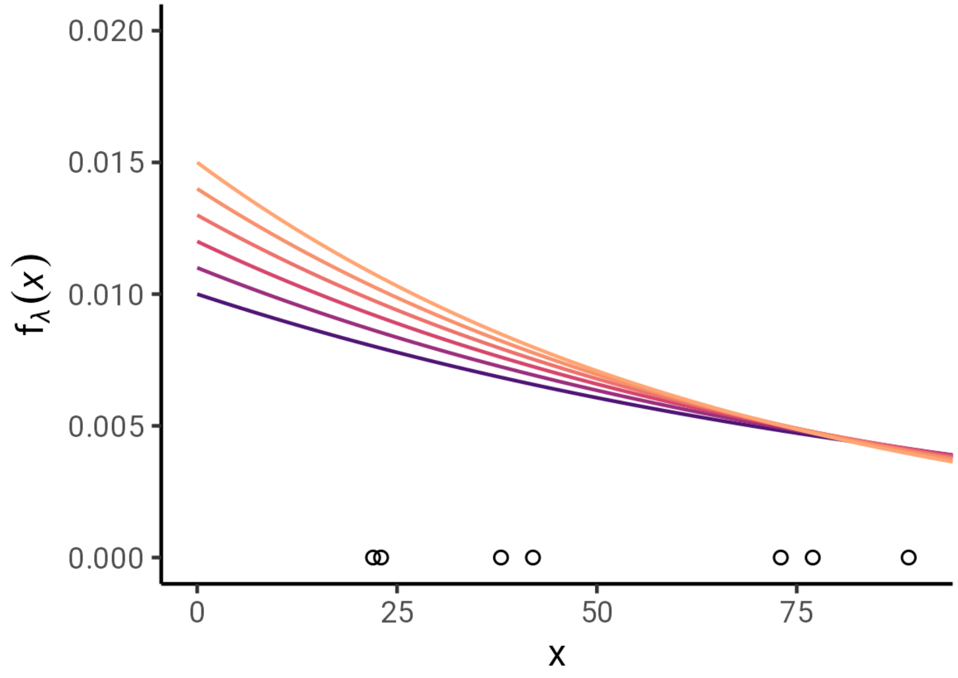


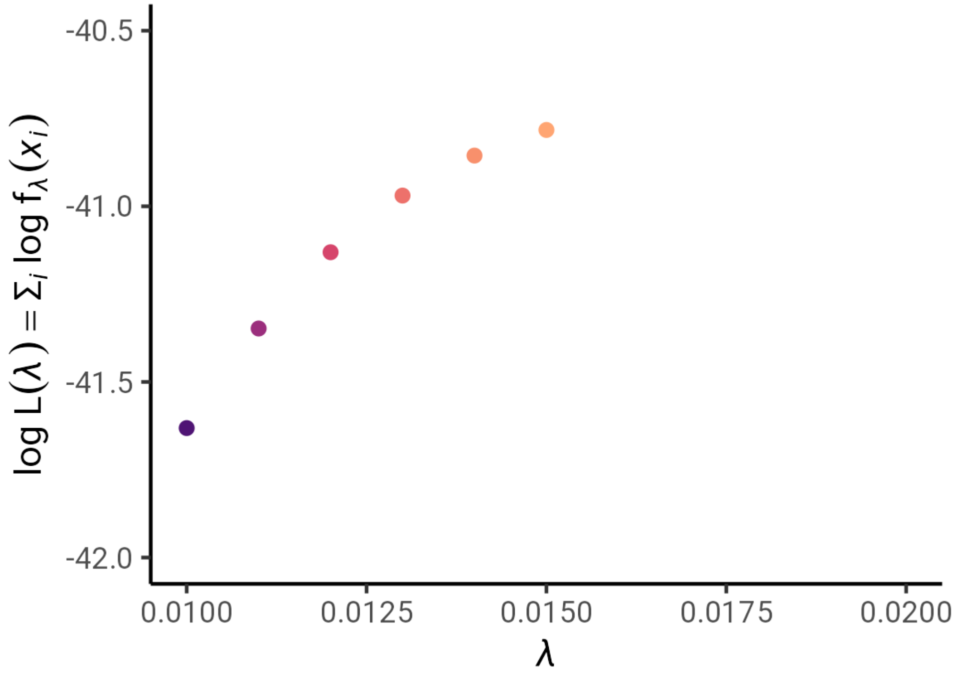


Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

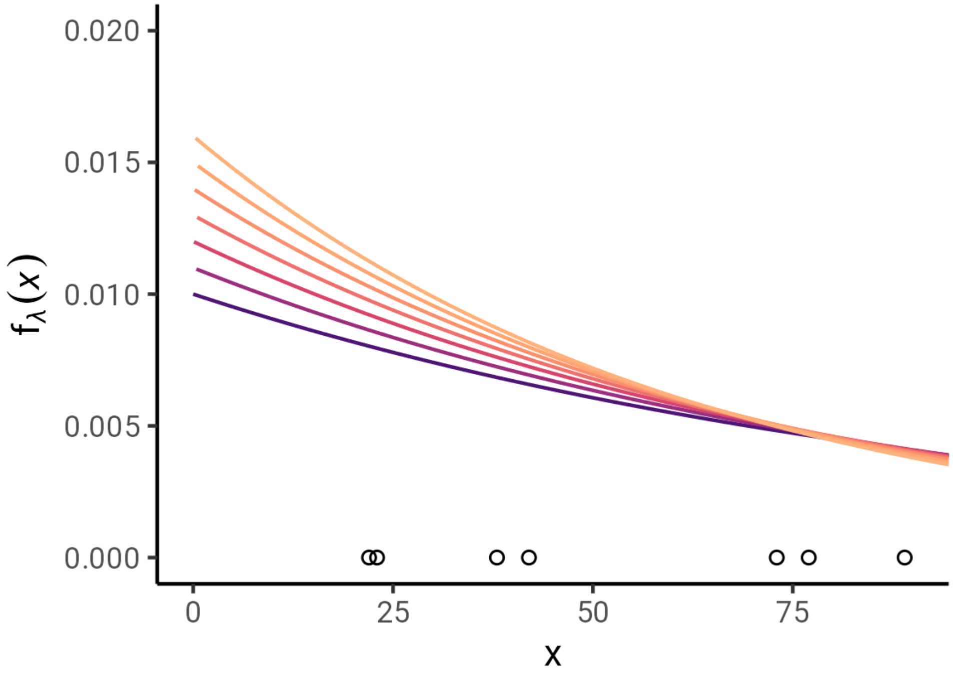
→ Maximum Likelihood Estimation

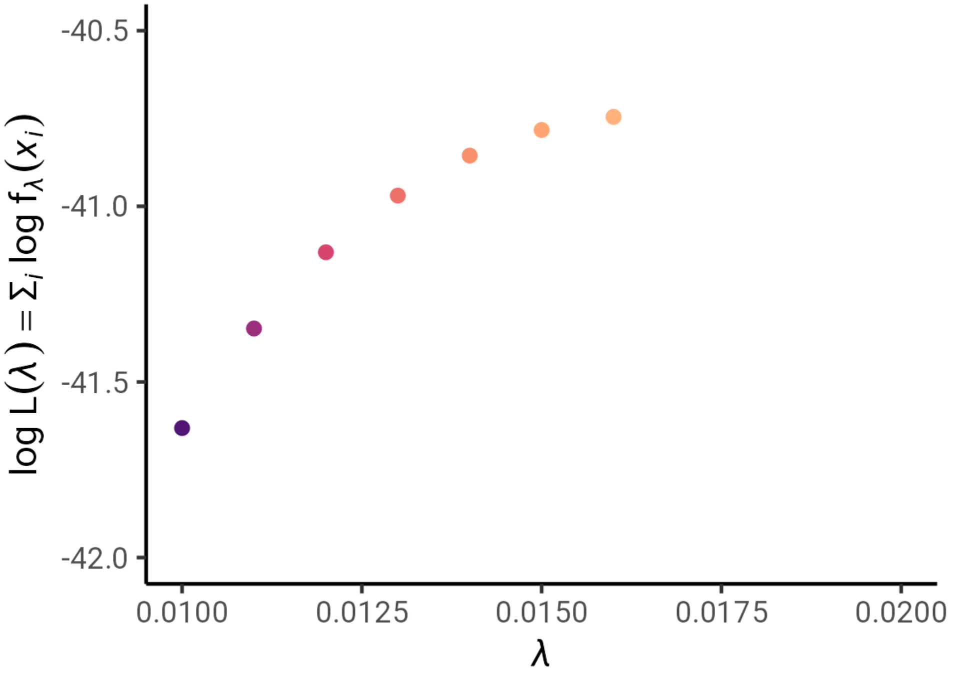




Inferring the Survival Density from Data

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

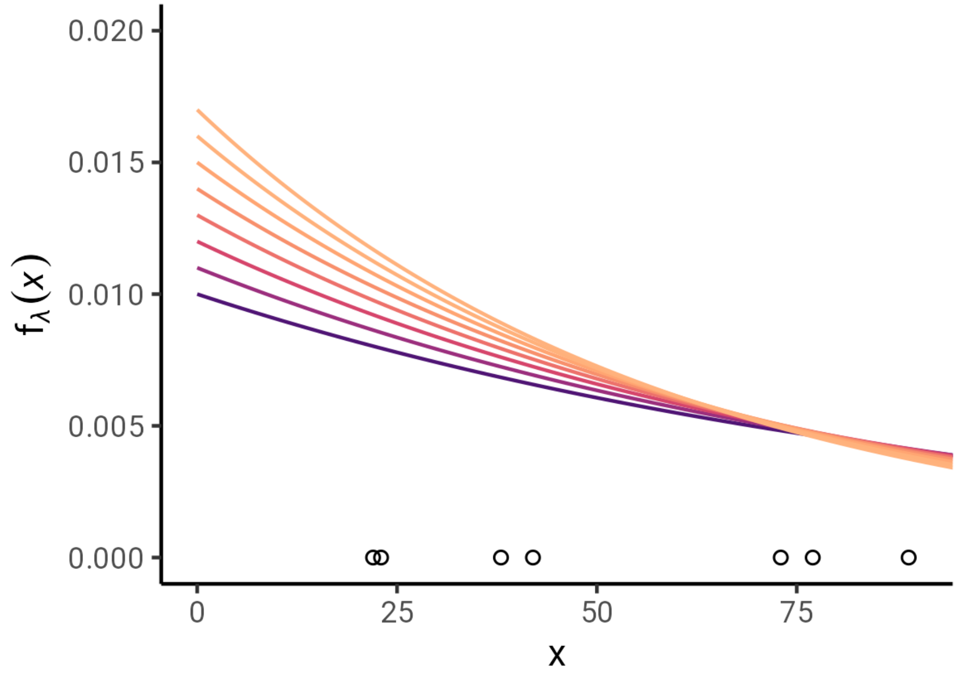


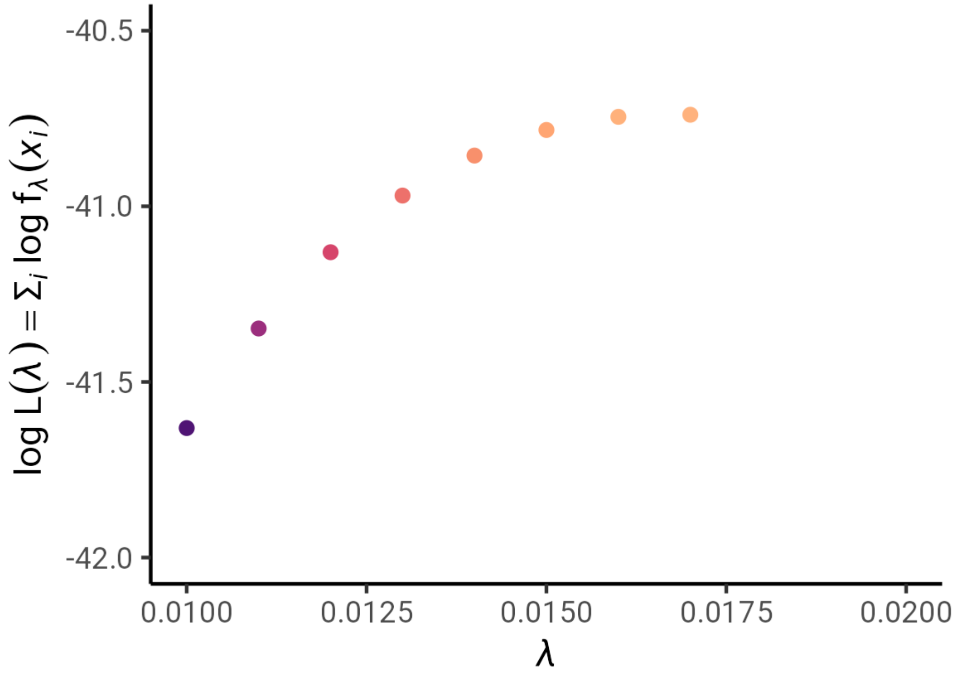


Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

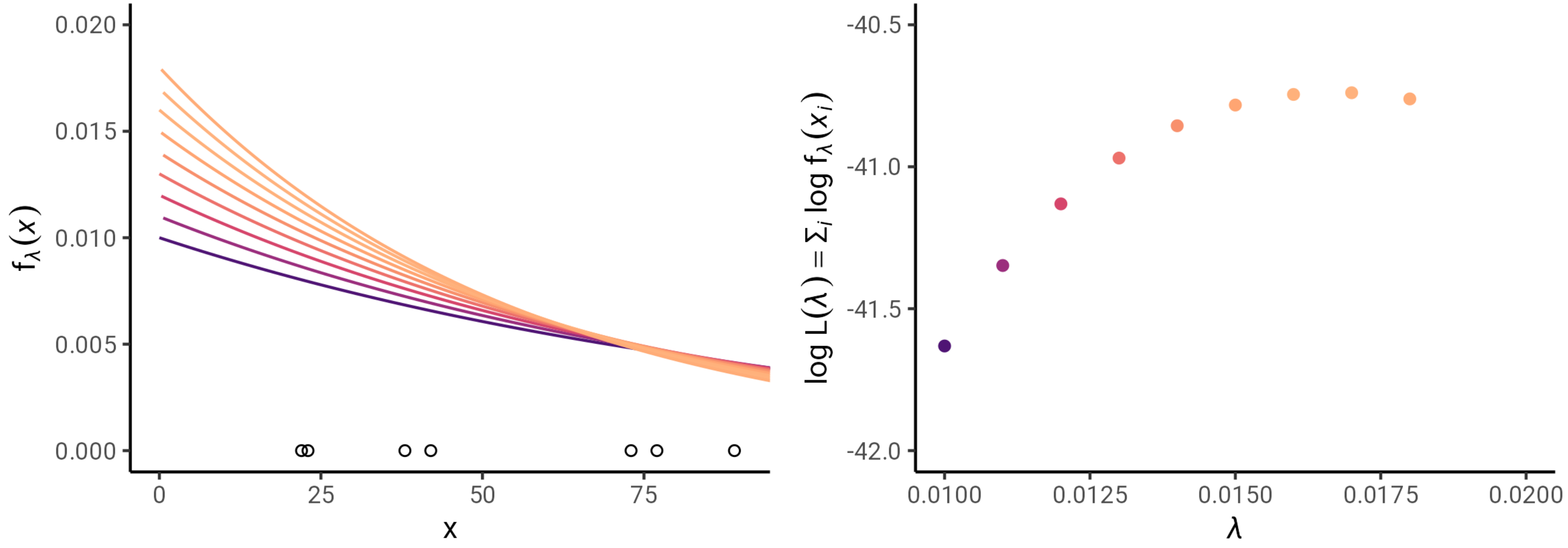




Inferring the Survival Density from Data

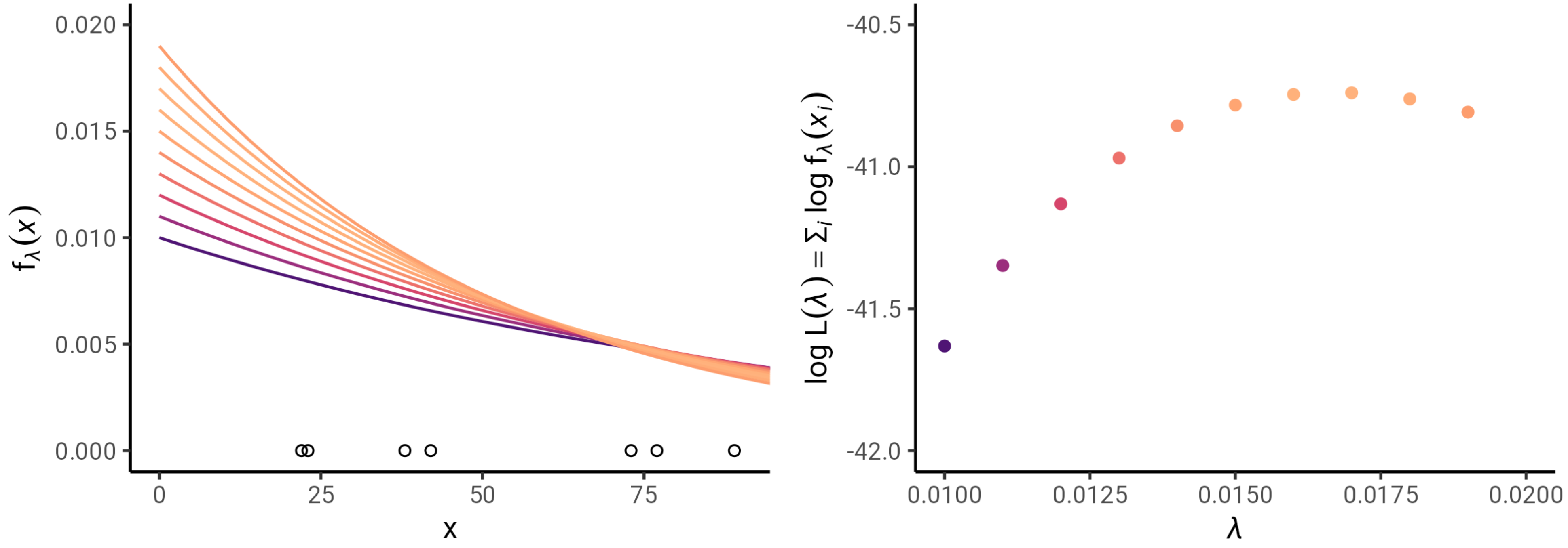
Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation



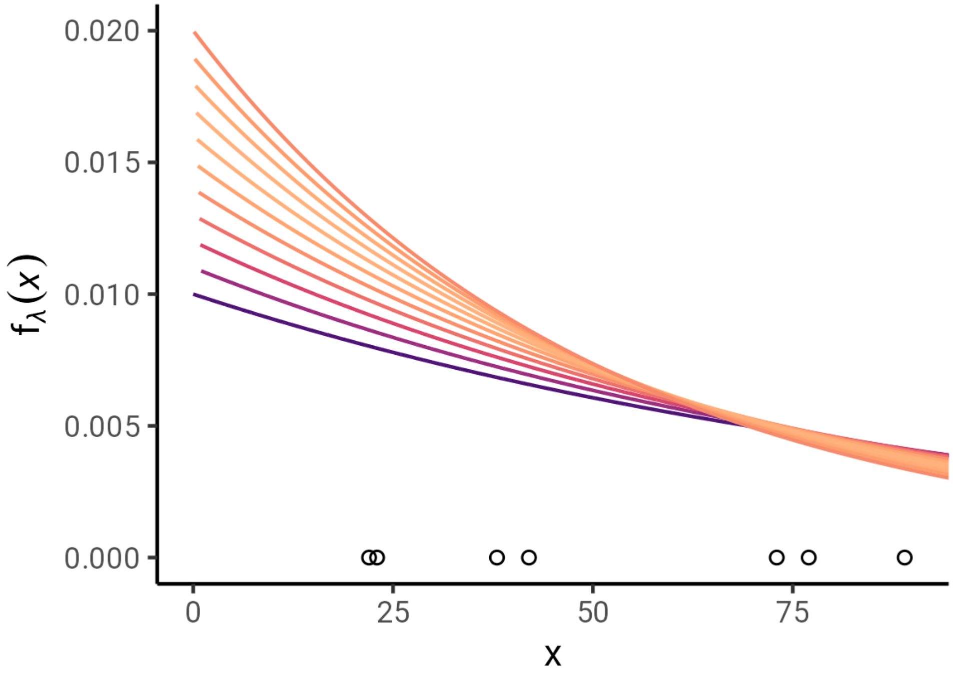
Inferring the Survival Density from Data

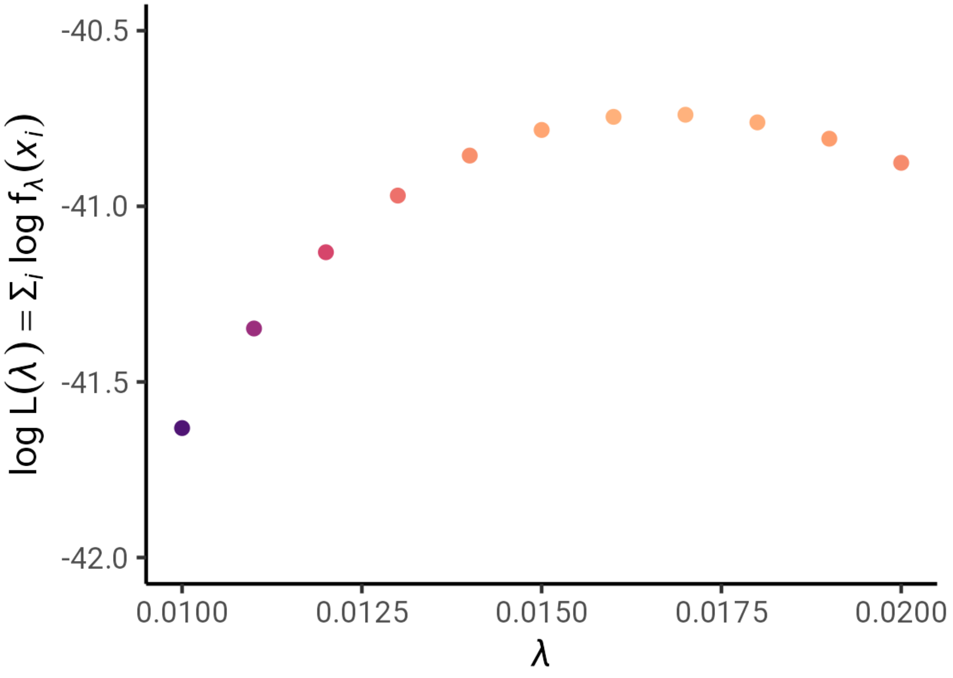
Which distribution most likely corresponds to the data?
→ Maximum Likelihood Estimation



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

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 $\boldsymbol{\theta}$

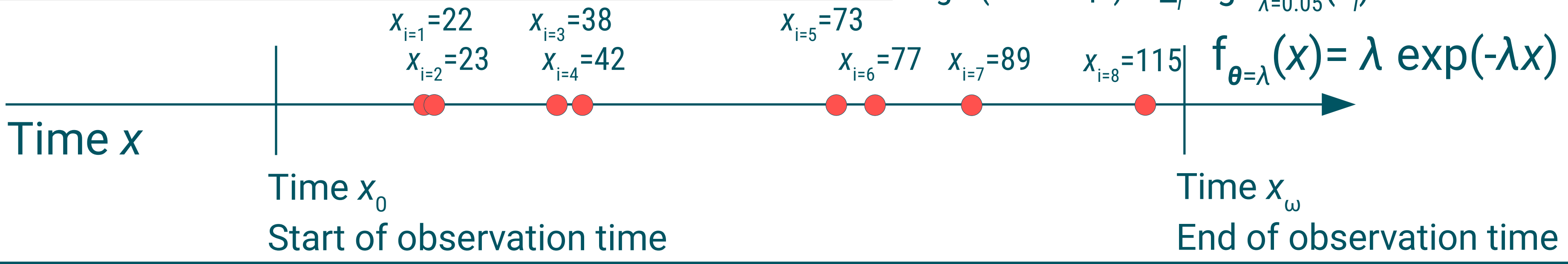
$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_i f_{\boldsymbol{\theta}}(x_i)$

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		
$\boldsymbol{i}$	$\boldsymbol{x}$	$f_{\lambda=0.05}(\boldsymbol{x})$	$\log f_{\lambda=0.05}(\boldsymbol{x})$
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?

→ Censored observations

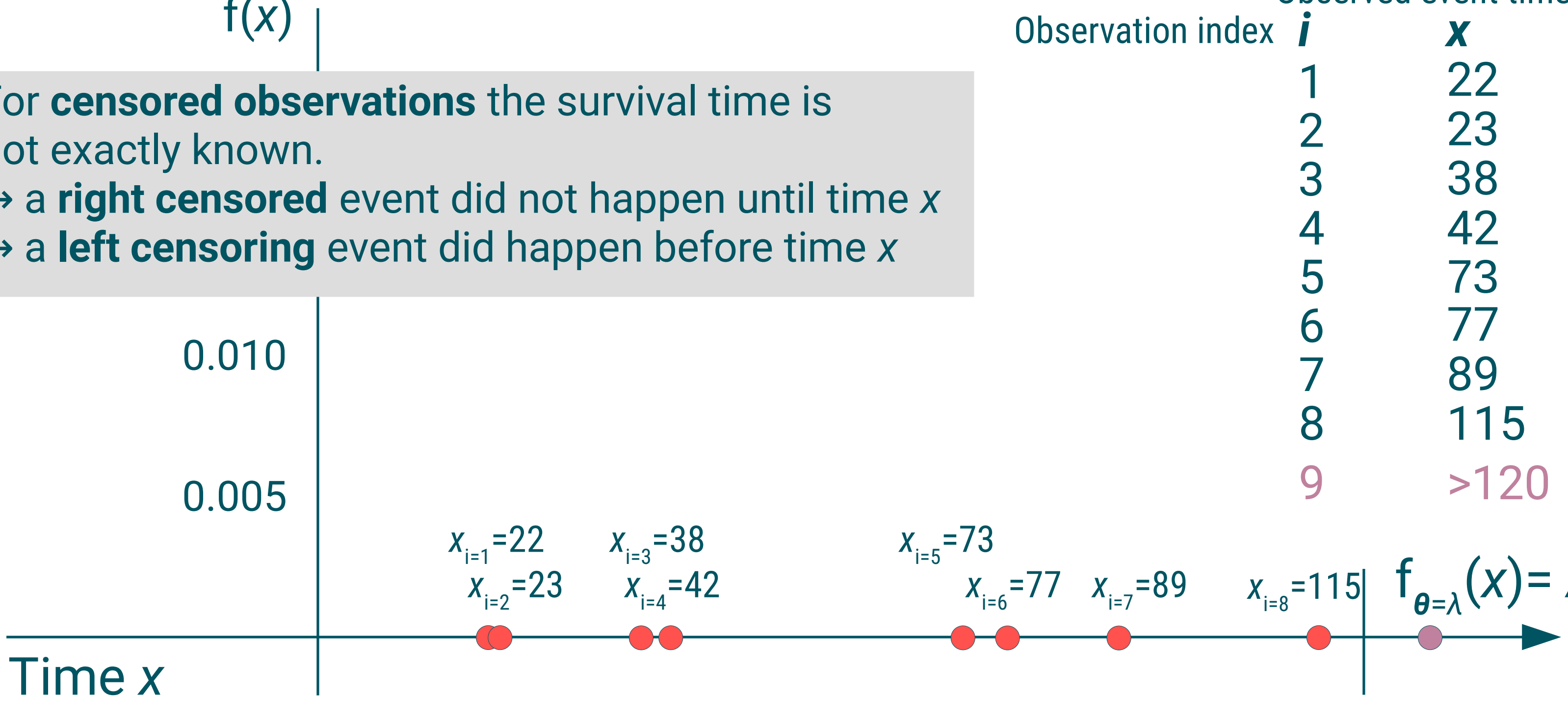
Inferring the Survival Density from Data

Which distribution most likely corresponds to the data?

→ Maximum Likelihood Estimation

For **censored observations** the survival time is not exactly known.
→ a **right censored** event did not happen until time x
→ a **left censoring** event did happen before time x

Observation index i	Observed event time x	Event indicator δ
1	22	1
2	23	1
3	38	1
4	42	1
5	73	1
6	77	1
7	89	1
8	115	1
9	>120	0



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

Time x_0
Start of observation time

Time x_ω
End of observation time

Inferring the Survival Density from Data

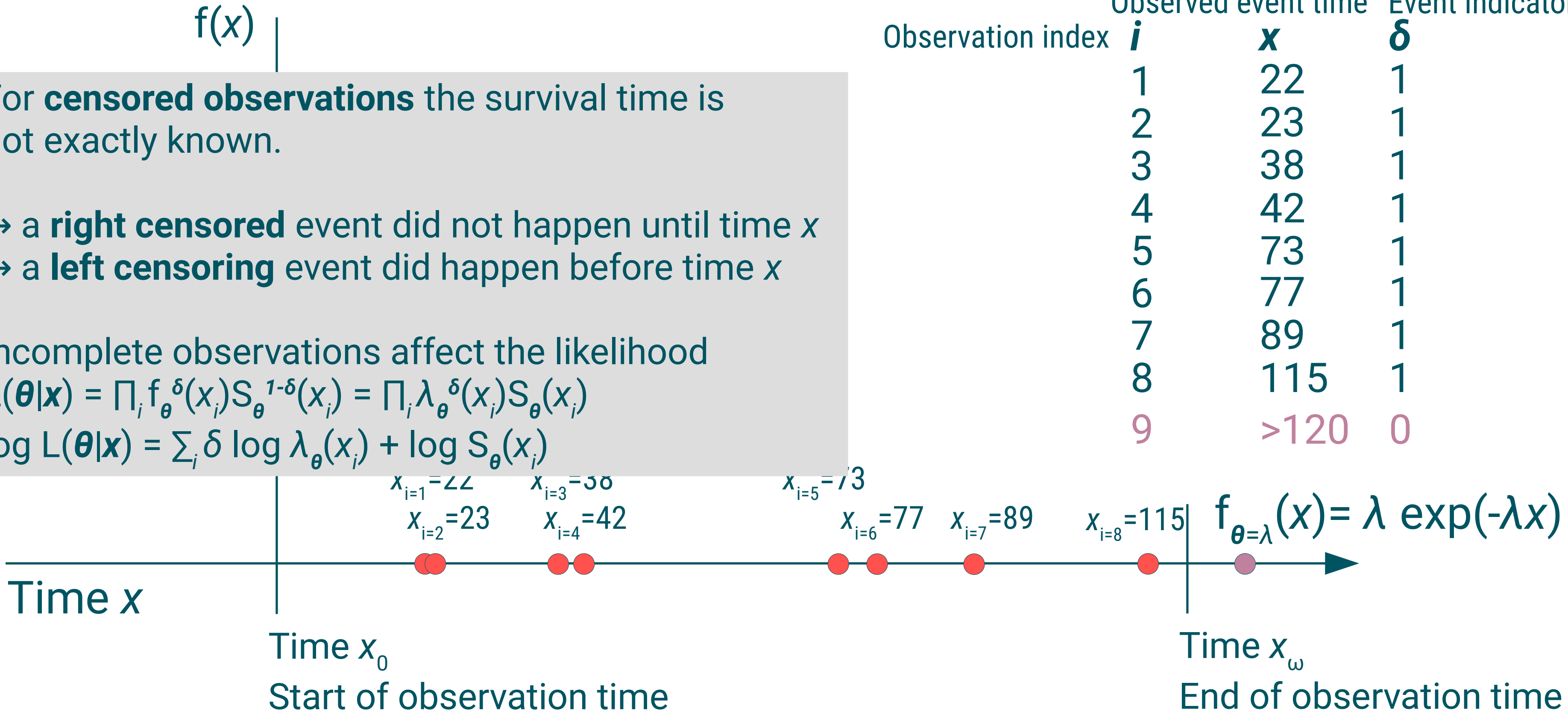
Which distribution most likely corresponds to the data?
→ Maximum Likelihood Estimation

For **censored observations** the survival time is not exactly known.

→ a **right censored** event did not happen until time x
→ a **left censoring** event did happen before time x

Incomplete observations affect the likelihood
$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_i f_{\boldsymbol{\theta}}^{\delta}(x_i) S_{\boldsymbol{\theta}}^{1-\delta}(x_i) = \prod_i \lambda_{\boldsymbol{\theta}}^{\delta}(x_i) S_{\boldsymbol{\theta}}(x_i)$
$\log L(\boldsymbol{\theta}|\boldsymbol{x}) = \sum_i \delta \log \lambda_{\boldsymbol{\theta}}(x_i) + \log S_{\boldsymbol{\theta}}(x_i)$

Observation index $\boldsymbol{i}$	Observed event time $\boldsymbol{x}$	Event indicator $\boldsymbol{\delta}$
1	22	1
2	23	1
3	38	1
4	42	1
5	73	1
6	77	1
7	89	1
8	115	1
9	>120	0



Materials for this lecture

github.com/jschoeley/survival_analysis-ur-ss22

Jonas Schöley

@jschoeley

0000-0002-3340-8518

j.schoeley@uni-rostock.de