Likelihood Function for Right Censored Data

Review

- Describe random variable for survival times
  - Survival function
  - Hazard function
  - Some common parametric models
- Censoring
  - Right censoring
  - Non-informative censoring

Likelihood function

- Likelihood function plays an important role in statistical estimation and inference.
- Let D denote the data, and $M(\theta)$ be a given model.
- For example,
  - D=$(T_1, \cdots, T_n)$: survival times on n individuals.
  - $M(\theta)$: $T_i \sim Exp(\theta)$ indepently,$i = 1, \cdots n$.
- The likelihood function describes the likelihood of observing the data D for a given model $M(\theta)$.

Likelihood function

- So we write
$$L(\theta) = L(D|M(\theta))$$
- It summarizes evidence about θ contained in D.
- Larger likelihood $\Leftrightarrow$Stronger evidence that data is generated from a model with that parameter.
- MLE: $\hat{\theta}$ maximize the likelihood function in the parameter space.

Likelihood: no censoring

- Suppose there is no censoring.
- Assume $\{T_1^0 \cdots, T_n^0\}$ independent and identically distributed with density $f_\theta(t)$.
- Then the likelihood function is given as
$$L(\theta) = \prod_{i=1}^{n} f_\theta(t_i^0)$$
where t_i^0 is the value that T_i^0 takes.
- Note that it is a function of θ, not $\{t_i^0\}$. (After the data is observed, $\{t_i^0\}$ are fixed.)

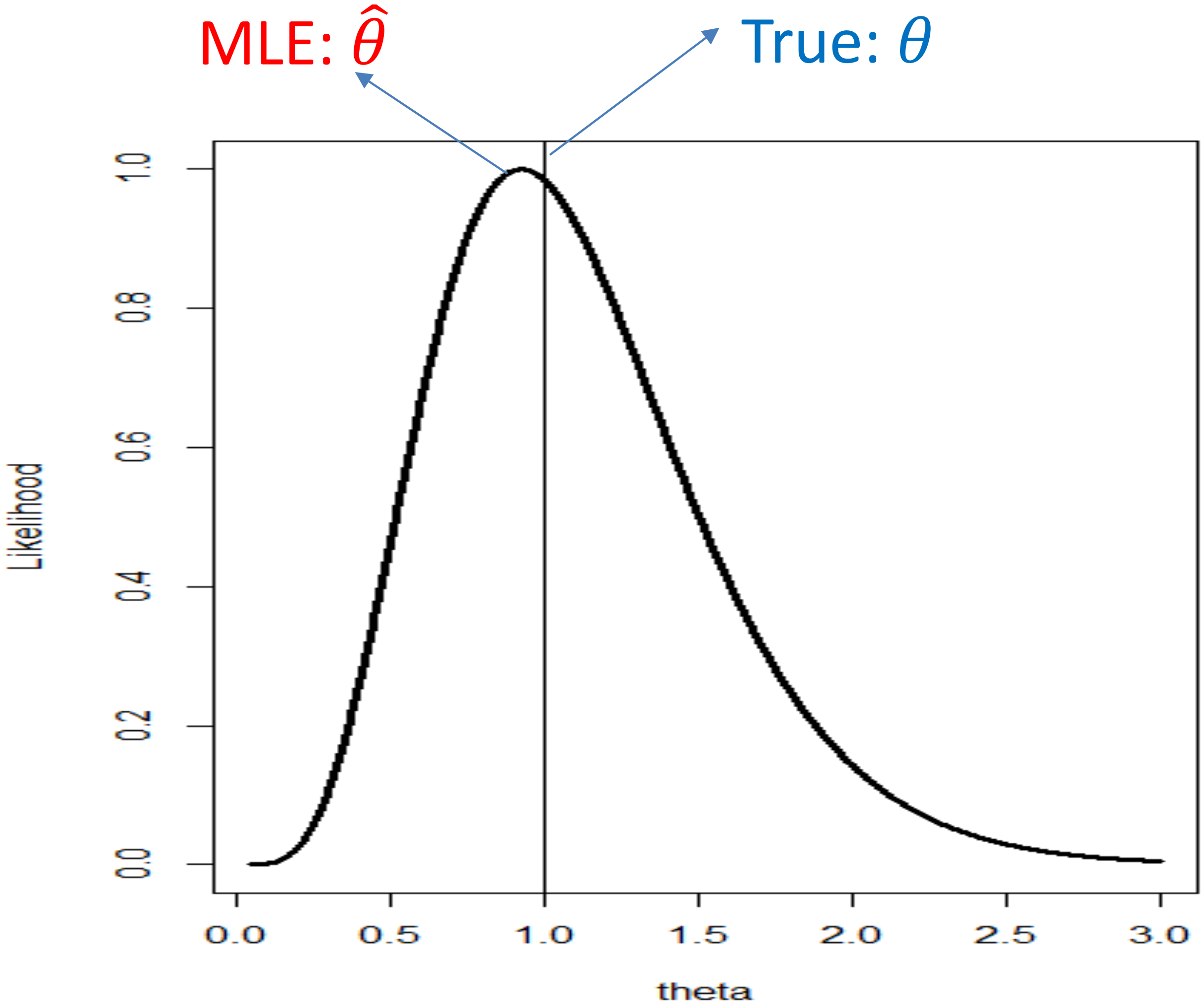
An example

- True survival times T_i^0 i.i.d.$\sim Exp(1)$.
- We observe the following data

  {0.1, 0.3, 0.7, 1.6, 2.7}.
- Suppose the truth that T_i^0 i.i.d.$\sim Exp(1)$ is unknown to the investigator.
- The investigator fits the data with $Exp(\theta)$.
- Then the likelihood function

$$L(\theta) = \prod_{i=1}^{5} \theta exp\,(-\theta t_i^0) = \theta^5 exp(-5.4\theta).$$

An example



MLE: $\hat{\theta}$
True: θ
Likelihood
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
theta

Likelihood function for censored data

- In survival analysis, some event times are observed, while others are only partially observed.
- We will show that the likelihood is well-defined for both cases.
- Likelihood function is a natural way to combine the information from fully and partially observed subjects.
- Likelihood function is the basis for all statistical methods in this class.

Likelihood for right-censored data

Let $(T_i, \delta_i)_{i=1}^{n}$ be a set of n independent observations from a right censored experiment.

- $T_i = \min(T_i^0, C_i)$, where T_i^0 is the true survival time, and C_i is the censoring time.
- $\delta_i = 1$, if an event occurred with $T_i = T_i^0$;
- $\delta_i = 0$, if censored with $T_i = C_i$.

Likelihood for right-censored data: Assumptions

We assume

- data is only subject to right censoring
- censoring is non-informative
- the observations from different subjects are independent and identically distributed.

Likelihood for right-censored data: An example

- Suppose true survival times $\{T_i^0\}_{i=1}^5$ i.i.d Exp(1)
- Suppose we observe {0.1, 0.3, 0.7, 1.6, 2.7}.
- Study ends at time t = 1 and censored if survival time longer than 1.
- We have censored data:

$$\{(0.1,1), (0.3,1), (0.7,1), (1,0), (1,0)\}.$$

Likelihood for right-censored data: An example

- Censored data:

$$\{(0.1,1),(0.3,1),(0.7,1),(1,0),(1,0)\}.$$

- For the first three fully observed data, the likelihood remains the same as before

$$L_1(\theta) = f_\theta(0.1) f_\theta(0.3) f_\theta(0.7)$$

- For two censored observations, the likelihood is

$$L_2(\theta) = P(T_4^0 > 1) P(T_5^0 > 1) = S_\theta(1) S_\theta(1).$$

- Combining the likelihoods together

$$L(\theta) = L_1(\theta) L_2(\theta) = \prod_{i=1}^{3} f_\theta(t_i) \prod_{i=4}^{5} S_\theta(t_i).$$

Summary

- Let $(t_i, \delta_i)_{i=1}^{n}$ be right censored data.
- The likelihood function is

$$L = \prod_{\text{uncensored}} f(t_i) \prod_{\text{Censored}} S(t_i)$$
$$= \prod_{i=1}^{n} [\, f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i}$$

Contribution of uncensored data

Contribution of right censored data

An example

- Recall for complete data: {0.1, 0.5, 0.5, 1.6, 2.7}.
- The likelihood function using complete data

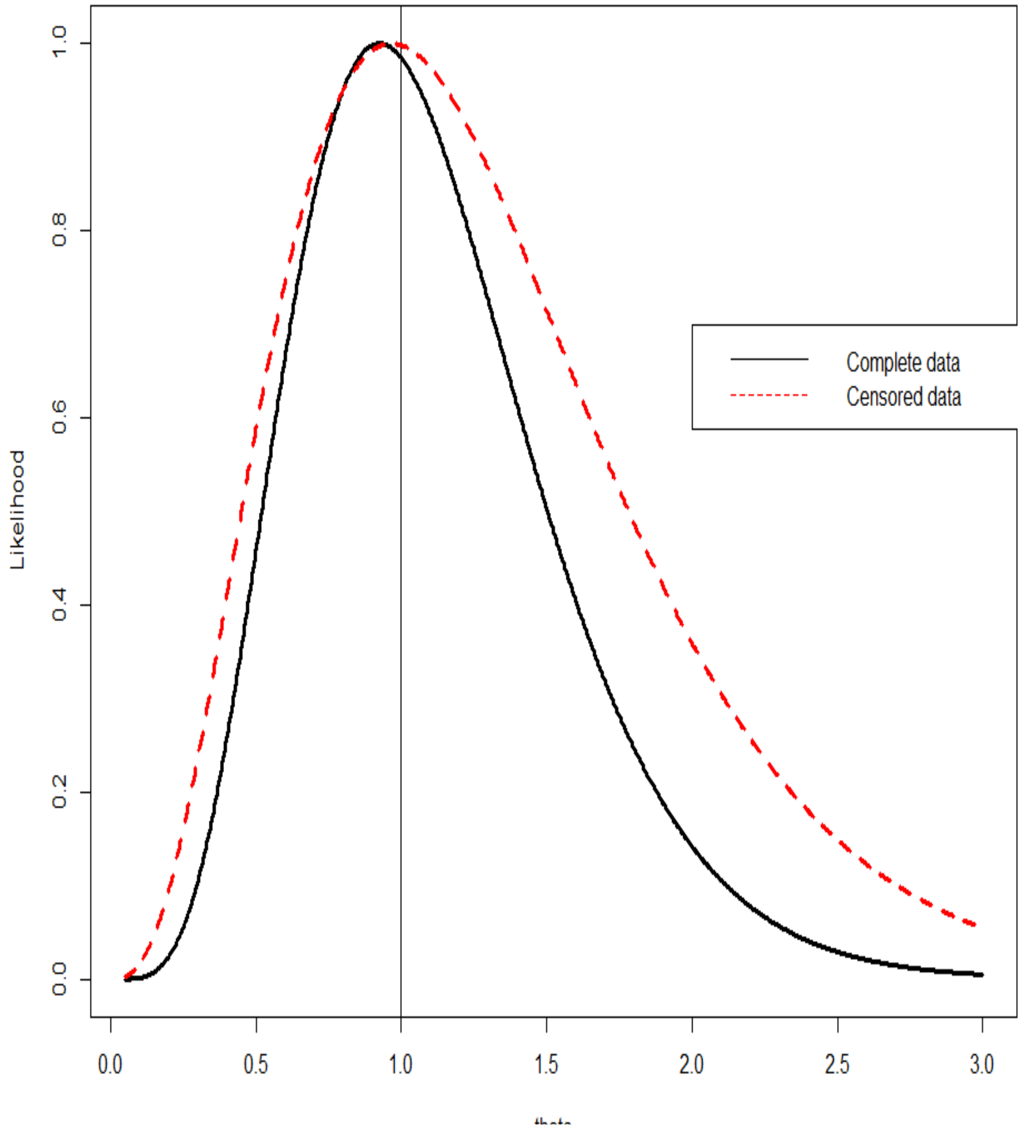
$$L_1(\theta) = \theta^5 exp(-\theta \sum_{i=1}^{5} t_i^0) = \theta^5 exp(-5.4\theta).$$

- The censored data: {0.1, 0.5, 0.5, 1, 1}.
- The likelihood function using censored data.

$$L_2(\theta) = f_\theta(0.1) f_\theta(0.3) f_\theta(0.7)\, f_\theta(0.1) S_\theta(1) S_\theta(1)$$

$$= \theta^3 exp(-3.1\theta)$$

An example



- Both seem to work well with the maximizers close to true parameter.
- The likelihood function for censored data is more wide.
- Censored data contains less information with wider range of likely values of θ.

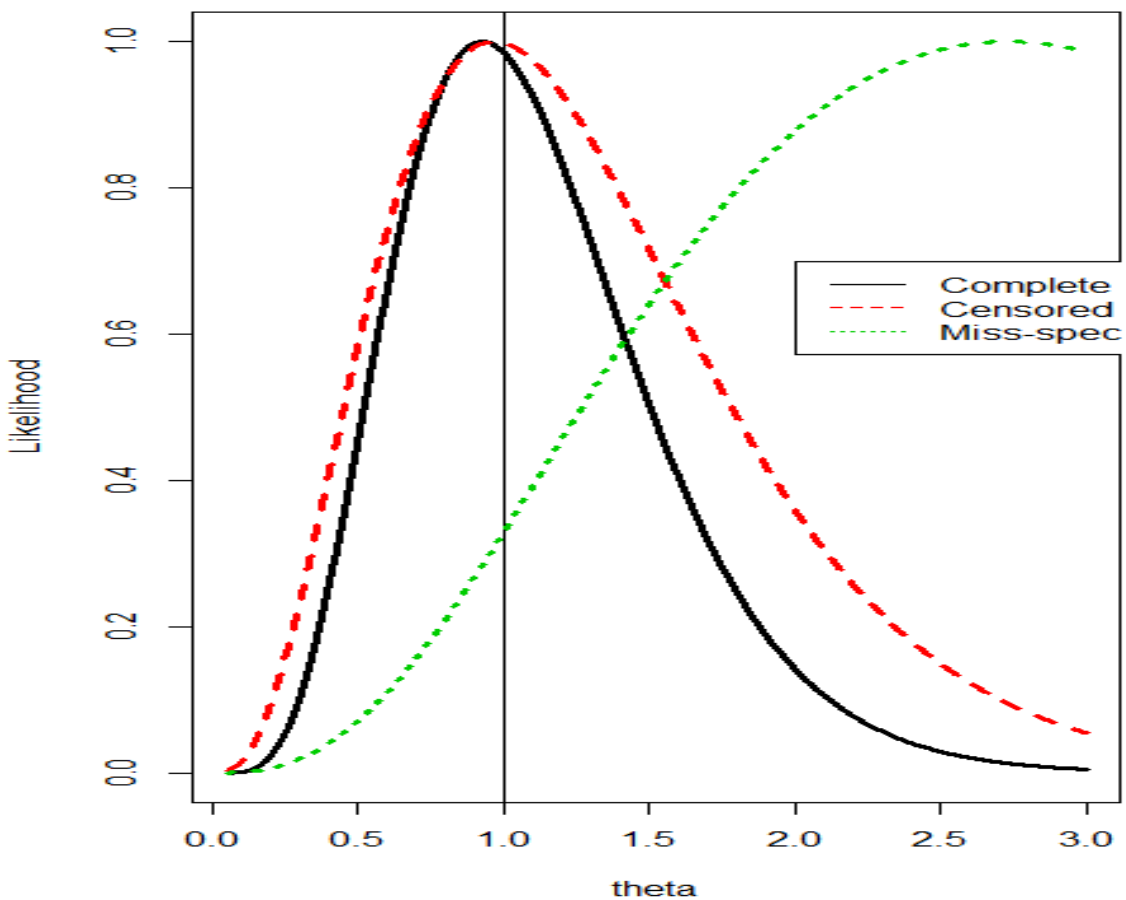
Effect of ignoring censoring

- If we ignore censoring, and plug censored data into the likelihood for complete data instead

$$L_3(\theta) = \prod_{i=1}^{n} \theta exp\,(-\theta t_i\ )=\theta^5 exp(-3.1\theta).$$

- It is a miss-specified likelihood function.

An example


Likelihood
0.0
0.2
0.4
0.6
0.8
1.0
Complete
Censored
Miss-spec
0.0
0.5
1.0
1.5
2.0
2.5
3.0
theta

Summary

- If we ignore the censoring in data. It might lead to biased conclusions.
- The likelihood function we constructed for censored data

$$L = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i}$$

- It appropriately combines information from both complete and censored data.
- It provides a basis for valid statistical analysis of right censored data.