Try MLE

2019-08-22

Hightlight

- To begin with an easy way, we may try the previous piecewise example, since it has already had close form for $S(t)$ and $f(t)$
- We set a true $\rho = 0.1$ at first and simulate data. However, the MLE of ρ does not at point 0.1. I think this may be becasue the formula of MLE is not correct (since it need independent assumption). We may need a now one?
- When write $\hat{S}(.)$ as the function of ρ, when we have true ρ, the $\hat{S}(.)$ is most close to the true ρ, which makes sense
- Using integral estimation method 2 is better than the Riemann integral estimation, since it does not need to estimate $\psi(t)$

Does the likelihood formula correct?

Previous we write likelihood as:

$$\prod_{i=1}^{n} f(\rho; t_i)^{\delta_i} S(\rho; t_i)^{1-\delta_i}$$

The log-likelihood is:

$$\sum_{i=1}^{n} \delta_i log(f(\rho; t_i)) + (1-\delta_i) S(\rho; t_i)$$

From the piecewise example, we have:

$S(\rho; t_i) = \frac{\rho-1}{\rho-2} exp(-2t) - \frac{1}{\rho-2} exp(-\rho t)$

$f(\rho; t_i) = \frac{2\rho-2}{\rho-2} exp(-2t) - \frac{\rho}{\rho-2} exp(-\rho t)$

$\frac{\partial f(\rho; t_i)}{\partial \rho} = exp(-2t_i)\big[-\frac{2}{(\rho-2)^2}\big] + exp(-\rho t_i)\big[\rho + \frac{2\rho}{\rho-2} + \frac{2}{(\rho-2)^2}\big]$

$\frac{\partial S(\rho; t_i)}{\partial \rho} = exp(-2t_i)\big[-\frac{1}{(\rho-2)^2}\big] + exp(-\rho t_i)\big[\frac{1}{(\rho-2)^2} + \frac{\rho}{\rho-2}\big]$

Then the derivation of loglikelihood re ρ is:

$$\begin{aligned}\frac{\partial l(\rho; t_i)}{\rho} =& \sum_{i=1}^{n} \delta_i \frac{exp(-2t_i)\big[-\frac{2}{(\rho-2)^2}\big] + exp(-\rho t_i)\big[\rho + \frac{2\rho}{\rho-2} + \frac{2}{(\rho-2)^2}\big]}{f(\rho; t_i)} \\ &+ \sum_{i=1}^{n} (1-\delta_i) \frac{exp(-2t_i)\big[-\frac{1}{(\rho-2)^2}\big] + exp(-\rho t_i)\big[\frac{1}{(\rho-2)^2} + \frac{\rho}{\rho-2}\big]}{S(\rho; t_i)}\end{aligned}$$

Then we may use Newton-Raphson method to estiamte its soluction. However.

Question: Does the formula correct? Is that special for independent scenarios? Since:

$$L(\theta; x, \delta) = \prod_{i=1}^{n} \big[f(x_i; \theta)\big]^{\delta_i} \big[S(x_i; \theta)\big]^{1-\delta_i}$$

This likelihood formula comes from:

$$L(\theta;x,\delta)=\prod_{i=1}^{n}f(x_i,\delta_i;\theta)=\prod_{i=1}^{n}f(x_i,\delta_i=1;\theta)^{\delta_i}f(x_i,\delta_i=0;\theta)^{1-\delta_i}$$

When $\theta_i=1$

$$\begin{aligned}P(x\le X<x+h,\delta_i=1)=&P(x\le X<x+h,C\ge T)\\ \approx&P(x\le T<x+h,C\ge x)\text{ (x is a fixed number and }h\to 0).\\ =&P(x\le T<x+h)P(C\ge x)\text{ (when independent)}\\ =&f(x)hP(C\ge x)\end{aligned}$$

Therefore,

$$f(x\le X<x+h,\delta_i=1)=\lim_{h\to 0}\frac{P(x\le T<x+h,\delta_i=1)}{h}=f(x)P(C\ge x)$$

Similarly, when $\delta_i=0$

$$\begin{aligned}P(x\le X<x+h,\delta_i=0)=&P(x\le X<x+h,T>C)\\ \approx&P(x\le C<x+h,T\ge x)\\ =&P(x\le C<x+h)P(T\ge x)\text{ (by independence)}\\ =&g(x)hP(T\ge x)\end{aligned}$$

where $g(x)$ is the pdf for censor time. Therefore,

$$f(x\le X<x+h,\delta_i=0)=\lim_{h\to 0}\frac{P(x\le C<x+h,\delta_i=0)}{h}=g(x)S(x)$$

Therefore,

$$\begin{aligned}f(x,\delta;\theta)=&\big[f(x)P(C\ge x)\big]^{\delta_i}\big[g(x)S(x)\big]^{1-\delta_i}\\ =&f(x)^{\delta_i}S(x)^{1-\delta_i}P(C\ge x)^{\delta_i}g(x)^{1-\delta_i}\end{aligned}$$

If only the θ is only the parameter for survival functions instead of censoring functions, then the $P(C\ge x)$ and $g(x)$ doesn't contain θ. The likelihood can be

$$L(\theta;x,\delta)=\prod_{i=1}^{n}\big[f(x_i;\theta)\big]^{\delta_i}\big[S(x_i;\theta)\big]^{1-\delta_i}$$

Therefore, this formula comes from the assumption of independence. (Did I think it in a right way?)

We may need a new formula for the dependent likelihood.

Try this likelihood function

Simulate data with size 1000, true $\rho=0.1$. The MLE function w.r.t ρ is

$$\begin{aligned}MLE(\rho)=&\sum_{i=1}^{n}\delta_i\frac{exp(-2t_i)\big[-\frac{2}{(\rho-2)^2}\big]+exp(-\rho t_i)\big[\rho+\frac{2\rho}{\rho-2}+\frac{2}{(\rho-2)^2}\big]}{f(\rho;t_i)}\\ &+\sum_{i=1}^{n}(1-\delta_i)\frac{exp(-2t_i)\big[-\frac{1}{(\rho-2)^2}\big]+exp(-\rho t_i)\big[\frac{1}{(\rho-2)^2}+\frac{\rho}{\rho-2}\big]}{S(\rho;t_i)}\end{aligned}$$

We could input the data to check who the MLE function value varies when ρ value changes.

I just used grid search, ρ from 0.1, 0.2, 0.3,. . . , 3.0. The corresponding MLE values:

```
##  [1]    768.695705    579.741892    452.814864    343.342147    237.089621
##  [6]    126.860701      7.549951   -125.603594   -277.984994   -456.476977
## [11]   -670.604278   -934.373854  -1269.598977  -1712.488837  -2328.105883
## [16]  -3246.479773  -4770.925365  -7811.292327 -16916.684274  19462.545076
## [21]  10358.229676   7319.662752   5797.750552   4882.659733   4271.101150
## [26]   3833.084702   3503.608046   3246.552840   3040.243220
```

The result that is most close to 0 is $\rho = 0.7$, which is not our true ρ value.

$\hat{S}(.)$ as ρ's function

When we write $\hat{S}(.)$ as a function of ρ, when the ρ is our true value, it works the best. And the integral estimation method that does not need to estimate the $\psi(t)$ works better than the Reimann integral estimation