Accelerated Failure Time

avinashbarnwal123

May 2019

1 Error - Normal Distribution

Assume the data follow below model:-

$$\begin{aligned} lny_i &= x_i{'}\beta + z_i\sigma \\ ln\hat{y}_i &= x_i{'}\hat{\beta} \end{aligned} \tag{1}$$

where y_i is the uncensored response and $\hat{y}_i$ is the predicted value.

1.1 Uncensored Data

1.1.1 Loss Function

Loss $function_i$ = -log-lik_i = -$log(f_{\hat{Y}_i}(\hat{y}_i))$
where, $f_{\hat{Y}}(\hat{y})$ is the probability density function(pdf) of $\hat{y}$.

$$\begin{aligned} z_i &= \frac{lny_i - ln\hat{y}_i}{\sigma} \sim f \\ f(z) &= \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \end{aligned} \tag{2}$$

Now using change of variable for probability density function(pdf). We can write the below equations.

$$\begin{aligned} f_{\hat{Y}}(\hat{y}) &= f_Z(z)\frac{\partial z}{\partial y} \\ f_{\hat{Y}}(\hat{y}) &= \frac{exp^{(\frac{(ln(z)-ln(\hat{y}))^2}{2\sigma^2}}}{\hat{y}\sigma\sqrt{2\pi}} \\ Lossfunction_i &= -log(\frac{exp^{(\frac{(ln(z_i)-ln(\hat{y}_i))^2}{2\sigma^2}}}{\hat{y}_i\sigma\sqrt{2\pi}}) \end{aligned} \tag{3}$$

1.1.2 Negative Gradient

In Gradient Boosting and Xgboost, we need to calculate negative gradient of loss function with respect to fitted value which is $\hat{y}$. Here i am changing between $\hat{y}_i$ to $\hat{y}$ to make it general.

$-\frac{\partial Lossfunction_i}{\partial \hat{y_i}} = \frac{\partial log(f_{\hat{Y_i}}(\hat{y_i}))}{\partial \hat{y_i}} = \frac{1}{f_{\hat{Y_i}}(\hat{y_i})} * \frac{\partial f_{\hat{Y_i}}(\hat{y_i})}{\partial \hat{y_i}}$

Using change of variable between z and $\hat{y}$. We can write the below in terms of z.

$$\frac{\partial f_{\hat{Y}}(\hat{y})}{\partial \hat{y}} = \frac{\partial}{\partial \hat{y}}[f_Z(z)\frac{\partial z}{\partial \hat{y}}] \tag{4}$$

Using product rule of differentiation, we can split $f_Z(z)$ and $\frac{\partial z}{\partial \hat{y}}$

$$= \frac{\partial f_Z(z)}{\partial \hat{y}}\frac{\partial z}{\partial \hat{y}} + f_Z(z)\frac{\partial^2 z}{\partial \hat{y}^2} \tag{5}$$

$$= \frac{\partial f_Z(z)}{\partial z}\frac{\partial z^2}{\partial \hat{y}^2} + f_Z(z)\frac{\partial^2 z}{\partial \hat{y}^2} \tag{6}$$

As we know $\frac{\partial f_Z(z)}{\partial z} = -zf_Z(z)$

Now, we will calculate the gradient of z with respect to $\hat{y}$ as we have $\frac{\partial z^2}{\partial \hat{y}^2}$ in the equation 6.

$$\frac{\partial z}{\partial \hat{y}} = \frac{-1}{\sigma \hat{y}} \tag{7}$$

We also need double differentiation of z with respect to $\hat{y}$ as we have $\frac{\partial^2 z}{\partial \hat{y}^2}$ in the equation 6.

$$\frac{\partial^2 z}{\partial \hat{y}^2} = \frac{1}{\sigma \hat{y}^2} \tag{8}$$

$$= \frac{-zf_Z(z)}{\sigma^2 \hat{y}^2} + \frac{f_Z(z)}{\sigma \hat{y}^2} \tag{9}$$

Now going back to original equation of calculating negative gradient of loss function with respect to $\hat{y_i}$. In equation 9, we have calculated the second part of negative gradient of loss function with respect to $\hat{y_i}$ which is $\frac{\partial f_{\hat{Y_i}}(\hat{y_i})}{\partial \hat{y_i}}$. By replacing the 9th equation in the negative gradient of loss function w.r.t. to $\hat{y_i}$, we get the results as follows:-

$$\begin{aligned} -\frac{\partial Lossfunction_i}{\partial \hat{y_i}} &= \frac{1}{f_{\hat{Y_i}}(\hat{y_i})} * \frac{\partial f_{\hat{Y_i}}(\hat{y_i})}{\partial \hat{y_i}} \\ &= \frac{1}{f_{\hat{Y_i}}(\hat{y_i})} * \{\frac{-z_i f_{Z_i}(z_i)}{\sigma^2 \hat{y_i}^2} + \frac{f_{Z_i}(z_i)}{\sigma \hat{y_i}^2}\} \\ &= \frac{f_{Z_i}(z_i)}{f_{\hat{Y_i}}(\hat{y_i})} * \{\frac{-z_i}{\sigma^2 \hat{y_i}^2} + \frac{1}{\sigma \hat{y}^2}\} \end{aligned} \tag{10}$$

In the above equations, we have $f_{Z_i}(z_i)$ and $f_{\hat{Y_i}}(\hat{y_i})$. As we know about the pdfs of z_i and $\hat{y_i}$.

Here, we replace the $z_i = \frac{lny_i - ln\hat{y}_i}{\sigma}$

$$f_{Z_i}(z_i) = \frac{e^{-z_i^2/2}}{\sqrt{2\pi}} = \frac{e^{-\frac{(lny_i - ln\hat{y}_i)^2}{2*\sigma^2}}}{\sqrt{2\pi}} \tag{11}$$

And we know that using change of variable , $\hat{y}_i$ follows log-normal distribution.

$$f_{\hat{Y}_i}(\hat{y}_i) = \frac{e^{-\frac{(lny_i - ln\hat{y}_i)^2}{2*\sigma^2}}}{\sqrt{2\pi} * \sigma * \hat{y}_i} \tag{12}$$

Combining equation 11 and 12, we have following results.

$$\frac{f_{Z_i}(z_i)}{f_{\hat{Y}_i}(\hat{y}_i)} = \sigma * \hat{y}_i \tag{13}$$

Coming back to original equation and then simplified form of equation 10. Replacing the result of equation 13 in equation 10. We get following results.

$$\begin{aligned} -\frac{\partial Lossfunction_i}{\partial \hat{y}_i} &= \frac{f_{Z_i}(z_i)}{f_{\hat{Y}_i}(\hat{y}_i)} * \{\frac{-z_i}{\sigma^2 \hat{y}_i{}^2} + \frac{1}{\sigma \hat{y}^2}\} \\ &= \sigma * \hat{y}_i * \{\frac{-z_i}{\sigma^2 \hat{y}_i{}^2} + \frac{1}{\sigma \hat{y}^2}\} \end{aligned} \tag{14}$$

Further z_i can be replace in terms of $\hat{y}_i$

$$\begin{aligned} &= -\frac{lny_i - ln\hat{y}_i}{\hat{y}_i \sigma^2} + \frac{1}{\hat{y}_i} \\ &= \frac{\sigma^2 - ln\frac{y_i}{\hat{y}_i}}{\sigma^2 \hat{y}_i} \end{aligned} \tag{15}$$