Accelerated Failure Time

Avinash Barnwal

May 2019

1 Error - Normal Distribution

Assume the data follow below model:-

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

where y_i is the uncensored response and $\hat{y}_i$ is the predicted value for i-th observation and σ is the standard deviation of the error.

1.1 Uncensored Data

1.1.1 Loss Function

$$L_i = -\text{log lik}_i = -log(f_{Y_i}(y_i|\hat{y}_i)) \tag{2}$$

where, $f_{\hat{Y}}(\hat{y})$ is the probability density function(pdf) of $\hat{y}$.

$$\begin{aligned} z_i &= \frac{\log y_i - \log \hat{y}_i}{\sigma} \sim f \\ f(z) &= \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \end{aligned} \tag{3}$$

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

$$\begin{aligned} f_Y(y|\hat{y}) &= f_Z(z)\frac{\partial z}{\partial y} \\ f_Y(y|\hat{y}) &= \frac{e^{(\frac{(ln(y)-ln(\hat{y}))^2}{2\sigma^2})}}{y\sigma\sqrt{2\pi}} \\ L_i &= -log(\frac{\exp(\frac{(ln(y_i)-ln(\hat{y}_i))^2}{2\sigma^2}}{y_i\sigma\sqrt{2\pi}}) \end{aligned} \tag{4}$$

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.

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

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

$$\frac{\partial f_Y(y|\hat{y})}{\partial \hat{y}} = \frac{\partial}{\partial \hat{y}}[f_Z(z)\frac{\partial z}{\partial \hat{y}}] \tag{5}$$

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

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

$$= \frac{\partial f_Z(z)}{\partial z}\frac{\partial z^2}{\partial \hat{y}^2} + f_Z(z)\frac{\partial^2 z}{\partial \hat{y}^2} \tag{7}$$

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 7.

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

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 7.

$$\frac{\partial^2 z}{\partial \hat{y}^2} = \frac{1}{\sigma \hat{y}^2} \tag{9}$$

$$= \frac{-zf_Z(z)}{\sigma^2 \hat{y}^2} + \frac{f_Z(z)}{\sigma \hat{y}^2} \tag{10}$$

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 $\frac{\partial f_{Y_i}(y_i|\hat{y_i})}{\partial \hat{y_i}}$. By replacing the 10th equation in the negative gradient of loss function w.r.t. to $\hat{y_i}$, we get the results as follows:-

$$\begin{aligned} -\frac{\partial L_i}{\partial \hat{y_i}} &= \frac{1}{f_{Y_i}(y_i|\hat{y_i})} * \frac{\partial f_{Y_i}(y_i|\hat{y_i})}{\partial \hat{y_i}} \\ = \frac{1}{f_{Y_i}(y_i|\hat{y_i})} &* \{\frac{-z_i f_{Z_i}(z_i)}{\sigma^2 \hat{y_i}^2} + \frac{f_{Z_i}(z_i)}{\sigma \hat{y_i}^2}\} \\ &= \frac{f_{Z_i}(z_i)}{f_{Y_i}(y_i|\hat{y_i})} * \{\frac{-z_i}{\sigma^2 \hat{y_i}^2} + \frac{1}{\sigma \hat{y}^2}\} \end{aligned} \tag{11}$$

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 $z_i = \frac{lny_i - ln\hat{y}_i}{\sigma}$

$$f_{Z_i}(z_i) = \frac{e^{-z_i^2/2}}{\sqrt{2\pi}} = \frac{e^{-\frac{(\log y_i - \log \hat{y}_i)^2}{2*\sigma^2}}}{\sqrt{2\pi}} \tag{12}$$

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

$$f_{Y_i}(y_i|\hat{y}_i) = \frac{e^{-\frac{(lny_i - ln\hat{y}_i)^2}{2*\sigma^2}}}{\sqrt{2\pi} * \sigma * y_i} \tag{13}$$

Combining equation 12 and 13, we have following results.

$$\frac{f_{Z_i}(z_i)}{f_{Y_i}(y_i|\hat{y}_i)} = \sigma * y_i \tag{14}$$

Coming back to original equation and then simplified form of equation 11. Replacing the result of equation 14 in equation 11. We get following results.

$$\begin{aligned} -\frac{\partial Lossfunction_i}{\partial \hat{y}_i} &= \frac{f_{Z_i}(z_i)}{f_{\hat{Y}_i}(\hat{y}_i)} * \{\frac{-z_i}{\sigma^2 \hat{y}_i{}^2} + \frac{1}{\sigma \hat{y}_i{}^2}\} \\ &= \sigma * y_i * \{\frac{-z_i}{\sigma^2 \hat{y}_i{}^2} + \frac{1}{\sigma \hat{y}_i{}^2}\} \end{aligned} \tag{15}$$

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

$$\begin{aligned} &= -\frac{(\log y_i - \log \hat{y}_i) * y_i}{\hat{y}_i{}^2 \sigma} + \frac{y_i}{\hat{y}_i{}^2} \\ &= \frac{y_i * \{\sigma - \log \frac{y_i}{\hat{y}_i}\}}{\sigma \hat{y}_i{}^2} \end{aligned} \tag{16}$$

1.1.3 Hessian

Hessian is the second derivative of Loss with respect to $\hat{y}_i$. As we have already calculated the negative gradient, further we need to take one more partial derivative with respect to $\hat{y}_i$ of negative gradient with negative sign.

$$\frac{\partial^2 L_i}{\partial \hat{y}_i{}^2} = \frac{\partial}{\partial \hat{y}_i} \frac{\partial L_i}{\partial \hat{y}_i} \tag{17}$$

After inputting the values of negative gradient calculated to above equation.

$$= \frac{\partial}{\partial \hat{y}_i} * \{\frac{y_i * \{-\sigma + \log \frac{y_i}{\hat{y}_i}\}}{\sigma \hat{y}_i{}^2}\} \tag{18}$$

$$= \frac{\partial}{\partial \hat{y}_i} * \{\frac{y_i * \{-\sigma + \log y_i - \log \hat{y}_i\}}{\sigma \hat{y}_i{}^2}\} \tag{19}$$

$$= \frac{\partial}{\partial \hat{y_i}} * \{ \frac{y_i * \{-\sigma + \log y_i\} - y_i * \log \hat{y_i}\}}{\sigma \hat{y_i}^2} \} \tag{20}$$

$$= \frac{-\hat{y_i}^2 * \frac{y_i}{\hat{y_i}} + 2 * y_i \hat{y_i} * \{+\sigma - \log \frac{y_i}{\hat{y_i}}\}}{\sigma \hat{y_i}^4} \tag{21}$$

To take derivative, we apply division rule of derivatives.

$$= \frac{y_i \hat{y_i} * \{2 * \sigma - 1 + \log \frac{y_i}{\hat{y_i}}\}}{\sigma \hat{y_i}^4} \tag{22}$$

$$= \frac{y_i \hat{y_i} * \{2 * \sigma - 1 + \log \frac{y_i}{\hat{y_i}}\}}{\sigma \hat{y_i}^4} \tag{23}$$

$$\frac{\partial^2 L_i}{\partial \hat{y_i}^2} = \frac{y_i \hat{y_i} * \{2 * \sigma - 1 + \log \frac{y_i}{\hat{y_i}}\}}{\sigma \hat{y_i}^4} \tag{24}$$

1.2 Left Censored Data

1.2.1 Loss Function

$$L_i = -\log \text{lik}_i = -log(F_{Y_i}(y_i|\hat{y_i})) \tag{25}$$

where, F is the cdf of $Y_i|\hat{y_i}$.

As we know,

$$log(Y_i) \sim N(log(\hat{y_i}), \sigma) \tag{26}$$

Therefore, equation 25 can be written as below using property of 26

$$L_i = -log(\phi\{\frac{log(\frac{y_i}{\hat{y_i}})}{\sigma}\}) \tag{27}$$

1.2.2 Negative Gradient

$$-\frac{\partial L_i}{\partial \hat{y_i}} = -\frac{\partial - log(\phi\{\frac{log(\frac{y_i}{\hat{y_i}})}{\sigma}\})}{\partial \hat{y_i}} \tag{28}$$

$$-\frac{\partial L_i}{\partial \hat{y_i}} = \frac{\partial log(\phi\{\frac{log(\frac{y_i}{\hat{y_i}})}{\sigma}\})}{\partial \hat{y_i}} \tag{29}$$

$$= \frac{\phi'\{\frac{log(\frac{y_i}{\hat{y_i}})}{\sigma}\}}{\phi\{\frac{log(\frac{y_i}{\hat{y_i}})}{\sigma}\}} \tag{30}$$

Now above $\phi'\{\frac{log(\frac{y_i}{\hat{y_i}})}{\sigma}\}$ is calculated first order derivative of cdf which pdf of normal distribution and we have inner $\frac{log(\frac{y_i}{\hat{y_i}})}{\sigma}$, where this is calculated based on chain rule.

$$\frac{\partial \frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}}{\partial \hat{y}_i} = \frac{-1}{\sigma \hat{y}_i} \tag{31}$$

Therefore, combining results of chain rule and pdf, below is the final result for negative gradient,

$$-\frac{\partial L_i}{\partial \hat{y}_i} = \frac{-f(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})}{\sigma \hat{y}_i \phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}} \tag{32}$$

1.2.3 Hessian

Hessian is the second derivative of Loss with respect to $\hat{y}_i$. As we have already calculated the negative gradient, further we need to take one more partial derivative with respect to $\hat{y}_i$ of negative gradient with negative sign.

$$\frac{\partial^2 L_i}{\partial \hat{y}_i{}^2} = \frac{\partial}{\partial \hat{y}_i} \frac{\partial L_i}{\partial \hat{y}_i} \tag{33}$$

After inputting the values of negative gradient calculated to above equation.

$$= \frac{\partial}{\partial \hat{y}_i} * \frac{f(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})}{\sigma \hat{y}_i \phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}} \tag{34}$$

As we know the derivative of f(z) is $-zf(z)$. Using this information for above equation , further this is reduce in below form. Here we also have $\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}$ in the inner variable , where we calculate the derivative of it using chain rule.

$$\frac{\partial \{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}}{\partial \hat{y}_i} = \frac{-1}{\sigma \hat{y}_i} \tag{35}$$

For equation 34, we apply division rule of derivative and we know

$$\frac{\partial \phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}}{\partial \hat{y}_i} = -\frac{f(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})}{\sigma \hat{y}_i} \tag{36}$$

$$\frac{\partial f\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}}{\partial \hat{y}_i} = \{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\} \frac{f(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})}{\sigma \hat{y}_i} \tag{37}$$

$$\frac{\partial^2 L_i}{\partial \hat{y}_i{}^2} = \frac{\hat{y}_i \phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\} f^{'}(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}) - f(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})\{\phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\} - \hat{y}_i f^{'}(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})\}}{\sigma \hat{y}_i{}^2 \phi^2\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}} \tag{38}$$

1.3 Right Censored Data

1.3.1 Loss Function

$$L_i = -\log\ \text{lik}_i = -log(1 - F_{Y_i}(y_i|\hat{y}_i)) \tag{39}$$

where, F is the cdf of $Y_i|\hat{y}_i$.

Therefore, equation 25 can be written as below using property of 26

$$L_i = -log(1 - \phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}) \tag{40}$$

1.3.2 Negative Gradient

$$-\frac{\partial L_i}{\partial \hat{y}_i} = -\frac{\partial - log(1 - \phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})}{\partial \hat{y}_i} \tag{41}$$

$$-\frac{\partial L_i}{\partial \hat{y}_i} = \frac{\partial log(1 - \phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})}{\partial \hat{y}_i} \tag{42}$$

Using similar steps, we have done for left censored.

$$-\frac{\partial L_i}{\partial \hat{y}_i} = -\frac{\phi'\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}}{\phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}} \tag{43}$$

which is nothing but negative of the negative gradient of left censored data. Therefore,

$$-\frac{\partial L_i}{\partial \hat{y}_i} = \frac{f(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})}{\sigma\hat{y}_i\phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}} \tag{44}$$

1.3.3 Hessian

Similar to left censored data, we have same value but with negative sign. Therefore,

$$\frac{\partial^2 L_i}{\partial \hat{y}_i{}^2} = -\frac{\hat{y}_i\phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}f'(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}) - f(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})\{\phi\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\} - \hat{y}_i f'(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})\}}{\sigma\hat{y}_i{}^2\phi^2\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\}} \tag{45}$$

1.4 Interval Censored Data

1.4.1 Loss Function

This is combination of left censored and right censored data.

$$L_i = -\log\ \text{lik}_i = -log(F_{Y_i^u}(y_i^u|\hat{y}_i) - F_{Y_i^l}(y_i^l|\hat{y}_i)) \tag{46}$$

where, F is the cdf of $Y_i|\hat{y}_i$, y_i^u is the upper limit of time and y_i^l is the lower limit of the time. Above equation is written in terms of ϕ notations as below.

$$L_i = -log(\phi\{\frac{log(\frac{y_i^u}{\hat{y}_i})}{\sigma}\} - \phi\{\frac{log(\frac{y_i^l}{\hat{y}_i})}{\sigma}\}) \tag{47}$$

1.4.2 Negative Gradient

$$-\frac{\partial L_i}{\partial \hat{y}_i} = -\frac{\partial - log(\phi\{\frac{log(\frac{y_i^u}{\hat{y}_i})}{\sigma}\} - \phi\{\frac{log(\frac{y_i^l}{\hat{y}_i})}{\sigma}\})}{\partial \hat{y}_i} \tag{48}$$

$$-\frac{\partial L_i}{\partial \hat{y}_i} = \frac{\phi^{'}\{\frac{log(\frac{y_i^u}{\hat{y}_i})}{\sigma}\} - \phi^{'}\{\frac{log(\frac{y_i^l}{\hat{y}_i})}{\sigma}\}}{\phi\{\frac{log(\frac{y_i^u}{\hat{y}_i})}{\sigma}\} - \phi\{\frac{log(\frac{y_i^l}{\hat{y}_i})}{\sigma}\}} \tag{49}$$

As we know from 36

$$\frac{\partial \phi\{log(\frac{y_i}{\hat{y}_i})\}}{\partial \hat{y}_i} = -\frac{f(\{\frac{log(\frac{y_i}{\hat{y}_i})}{\sigma}\})}{\sigma \hat{y}_i} \tag{50}$$

By replacing the above equation in 49, we get the results mentioned below:-

$$-\frac{\partial L_i}{\partial \hat{y}_i} = \frac{-\frac{f(\{\frac{log(\frac{y_i^u}{\hat{y}_i})}{\sigma}\})}{\sigma \hat{y}_i} + \frac{f(\{\frac{log(\frac{y_i^l}{\hat{y}_i})}{\sigma}\})}{\sigma \hat{y}_i}}{\phi\{\frac{log(\frac{y_i^u}{\hat{y}_i})}{\sigma}\} - \phi\{\frac{log(\frac{y_i^l}{\hat{y}_i})}{\sigma}\}} \tag{51}$$

$$-\frac{\partial L_i}{\partial \hat{y}_i} = \frac{-f(\{\frac{log(\frac{y_i^u}{\hat{y}_i})}{\sigma}\}) + f(\{\frac{log(\frac{y_i^l}{\hat{y}_i})}{\sigma}\})}{\sigma \hat{y}_i\{\phi\{\frac{log(\frac{y_i^u}{\hat{y}_i})}{\sigma}\} - \phi\{\frac{log(\frac{y_i^l}{\hat{y}_i})}{\sigma}\}\}} \tag{52}$$

1.4.3 Hessian

Hessian is the second derivative of Loss with respect to $\hat{y}_i$. As we have already calculated the negative gradient, further we need to take one more partial derivative with respect to $\hat{y}_i$ of negative gradient with negative sign.

$$\frac{\partial^2 L_i}{\partial {\hat{y}_i}^2} = \frac{\partial}{\partial \hat{y}_i}\frac{\partial L_i}{\partial \hat{y}_i} \tag{53}$$

After inputting the values of negative gradient calculated to above equation.

$$= \frac{\partial}{\partial \hat{y}_i} * \frac{f(\{\frac{log(\frac{y_i^u}{\hat{y}_i})}{\sigma}\}) - f(\{\frac{log(\frac{y_i^l}{\hat{y}_i})}{\sigma}\})}{\sigma \hat{y}_i\{\phi\{\frac{log(\frac{y_i^u}{\hat{y}_i})}{\sigma}\} - \phi\{\frac{log(\frac{y_i^l}{\hat{y}_i})}{\sigma}\}\}} \tag{54}$$

Using division rule using the properties of $f^{'}(z) = -zf(z)$, 37 and 38. Below is the new equation, we have split the above equation numerator into 2 parts.

$$= \frac{\hat{y_i}\{\phi\{\frac{log(\frac{y_i^u}{\hat{y_i}})}{\sigma}\} - \phi\{\frac{log(\frac{y_i^l}{\hat{y_i}})}{\sigma}\}\} * \{f^{'}(\{\frac{log(\frac{y_i^u}{\hat{y_i}})}{\sigma}\})\} - f(\{\frac{log(\frac{y_i^u}{\hat{y_i}})}{\sigma}\})\{\{\phi\{\frac{log(\frac{y_i^u}{\hat{y_i}})}{\sigma}\} - \phi\{\frac{log(\frac{y_i^l}{\hat{y_i}})}{\sigma}\}\} + \hat{y_i}\{\phi^{'}\{\frac{log(\frac{y_i^u}{\hat{y_i}})}{\sigma}\} - \phi^{'}\{\frac{log(\frac{y_i^l}{\hat{y_i}})}{\sigma}\}\}\}}{\sigma\hat{y_i}^2\{\phi\{\frac{log(\frac{y_i^u}{\hat{y_i}})}{\sigma}\} - \phi\{\frac{log(\frac{y_i^l}{\hat{y_i}})}{\sigma}\}\}^2} \tag{55}$$

Similarly we can do it for second part.

2 Error - Logistic Distribution

2.1 Uncensored Data

2.1.1 Loss Function

$$L_i = -\log \text{lik}_i = -log(f_{Y_i}(y_i|\hat{y_i})) \tag{56}$$

where, $f_{\hat{Y}}(\hat{y})$ is the probability density function(pdf) of $\hat{y}$.

$$\begin{aligned} z_i &= \frac{\log y_i - \log \hat{y_i}}{\sigma} \sim f \\ f(z) &= \frac{e^z}{(1+e^z)^2} \end{aligned} \tag{57}$$

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

$$\begin{aligned} f_Y(y|\hat{y}) &= f_Z(z)\frac{\partial z}{\partial y} \\ f_Y(y|\hat{y}) &= \frac{e^{(\frac{(ln(y)-ln(\hat{y}))}{\sigma})}}{\sigma * y * (1+e^{(\frac{(ln(y)-ln(\hat{y}))}{\sigma})})^2} \\ L_i &= -log(\frac{e^{(\frac{(ln(y)-ln(\hat{y}))}{\sigma})}}{\sigma * y * (1+e^{(\frac{(ln(y)-ln(\hat{y}))}{\sigma})})^2}) \\ &= -log(\frac{e^{(\frac{ln(\frac{y}{\hat{y}})}{\sigma})}}{\sigma * y * (1+e^{(\frac{(ln\frac{y}{\hat{y}})}{\sigma})})^2}) \end{aligned} \tag{58}$$

2.1.2 Negative Gradient

Key results

$$\begin{aligned} w &= e^z \\ f^{'}(z) &= f(z)\frac{1-w}{1+w} \end{aligned} \tag{59}$$

Therefore, $f'(y|\hat{y})$ is calculated based on chain rule.

$$\frac{\partial f_Y(y|\hat{y})}{\partial \hat{y}} = \frac{\partial}{\partial \hat{y}}[f_Z(z)\frac{\partial z}{\partial \hat{y}}] \tag{60}$$

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

$$= \frac{\partial f_Z(z)}{\partial \hat{y}}\frac{\partial z}{\partial \hat{y}} + f_Z(z)\frac{\partial^2 z}{\partial \hat{y}^2} \tag{61}$$

$$= \frac{\partial f_Z(z)}{\partial z}\frac{\partial z^2}{\partial \hat{y}^2} + f_Z(z)\frac{\partial^2 z}{\partial \hat{y}^2} \tag{62}$$

$$\frac{\partial z}{\partial \hat{y}} = \frac{-1}{\sigma \hat{y}} \tag{63}$$

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 63.

$$\frac{\partial^2 z}{\partial \hat{y}^2} = \frac{1}{\sigma \hat{y}^2} \tag{64}$$

$$\frac{\partial^2 z}{\partial \hat{y}^2} = \frac{1}{\sigma \hat{y}^2} \tag{65}$$

Therefore,

$$-f^{'}(y|\hat{y}) = -\frac{f_Z^{'}(z)}{\sigma^2 \hat{y}^2} - \frac{f_Z(z)}{\sigma \hat{y}^2} \tag{66}$$

$$-f^{'}(y_i|\hat{y}_i) = -\frac{f_{Z_i}^{'}(z_i)}{\sigma^2 \hat{y_i}^2} - \frac{f_{Z_i}(z_i)}{\sigma \hat{y_i}^2} \tag{67}$$

2.1.3 Hessian

$$\begin{aligned} f^{''}(y_i|\hat{y}_i) &= \frac{\partial}{\partial \hat{y}_i}\{\frac{f_{Z_i}^{'}(z_i)}{\sigma^2 \hat{y_i}^2} + \frac{f_{Z_i}(z_i)}{\sigma \hat{y_i}^2}\} \\ &= \frac{\partial}{\partial \hat{y}_i}\{\frac{f_{Z_i}^{'}(z_i)}{\sigma^2 \hat{y_i}^2}\} + \frac{\partial}{\partial \hat{y}_i}\{\frac{f_{Z_i}(z_i)}{\sigma \hat{y_i}^2}\}\} \end{aligned} \tag{68}$$

First part is

$$\begin{aligned} \frac{\partial}{\partial \hat{y}_i}\{\frac{f_{Z_i}^{'}(z_i)}{\sigma^2 \hat{y_i}^2}\} &= \frac{1}{\sigma^2 \hat{y_i}^2}\frac{\partial}{\partial \hat{y}_i} f_{Z_i}^{'}(z_i) + f_{Z_i}^{'}(z_i)\frac{\partial}{\partial \hat{y}_i}\frac{1}{\sigma^2 \hat{y_i}^2} \\ &= -\frac{f_{Z_i}^{''}(z_i)}{\sigma^3 \hat{y_i}^3} + \frac{-2 * f_{Z_i}^{'}(z_i)}{\sigma^2 \hat{y_i}^3} \end{aligned} \tag{69}$$

Second part is

$$\begin{aligned}\frac{\partial}{\partial \hat{y_i}}\{\frac{f_{Z_i}(z_i)}{\sigma \hat{y_i}^2}\} &= \frac{1}{\sigma \hat{y_i}^2}\frac{\partial}{\partial \hat{y_i}} f_{Z_i}(z_i) + f_{Z_i}(z_i)\frac{\partial}{\partial \hat{y_i}}\frac{1}{\sigma \hat{y_i}^2} \\ &= -\frac{f'_{Z_i}(z_i)}{\sigma^2 \hat{y_i}^3} - \frac{2 * f_{Z_i}(z_i)}{\sigma \hat{y_i}^3}\end{aligned} \tag{70}$$

2.2 Left Censored Data

2.2.1 Loss Function

$$L_i = -\text{log lik}_i = -log(F_{Y_i}(y_i|\hat{y_i})) \tag{71}$$

where, F is the cdf of $Y_i|\hat{y_i}$.

As we know,

$$log(Y_i) \sim Logistic(log(\hat{y_i}), \sigma) \tag{72}$$

Therefore, equation 25 can be written as below using property of 26

$$L_i = -log(\phi\{\frac{log(\frac{y_i}{\hat{y_i}})}{\sigma}\}) \tag{73}$$

2.2.2 Negative Gradient

2.2.3 Hessian