Definition
$\tilde{C}_i$: True censoring indicator ($\tilde{C}_i = 0$ if censored, $\tilde{C}_i = 1$ otherwise)
C_i : Observed censoring indicator ($C_i = 0$ if censored, $C_i = 1$ otherwise)

Model setting
α_i : Probability that the observation i is truly censored.

$$\alpha_i = P(\tilde{C}_i = 0) \tag{4}$$

Then unconditional density(likelihood) for the observations that has $C_i = 1$ is as follows

$$\begin{aligned} P(T_i = t_i, C_i = 1) =& P(\tilde{C}_i = 0)P(T_i = t_i, C_i = 1|\tilde{C}_i = 0) \\ &+ P(\tilde{C}_i = 1)P(T_i = t_i, C_i = 1|\tilde{C}_i = 1) \\ =& \alpha_i S(t_i) + (1 - \alpha_i) f(t_i) \end{aligned} \tag{5}$$

And unconditional density(likelihood) for the observations that has $C_i = 0$ is as follows

$$\begin{aligned} P(T_i = t_i, C_i = 0) =& P(\tilde{C}_i = 0)P(T_i = t_i, C_i = 0|\tilde{C}_i = 0) \\ &+ P(\tilde{C}_i = 1)P(T_i = t_i, C_i = 0|\tilde{C}_i = 1) \\ =& \alpha_i S(t_i) \end{aligned} \tag{6}$$

(in (6), $P(T_i = t_i, C_i = 0|\tilde{C}_i = 1) = 0$ since an observation for which the event happens before the end of observing time can not be observed as censored.)

Hence, the likelihood for our model is as follows.

$$L(t_i) = \prod_{i=1}^{N} [\alpha_i S(t_i|X, \beta) + (1 - \alpha_i) f(t_i|X, \beta)]^{C_i} [\alpha_i S(t_i|X, \beta)]^{1-C_i} \tag{7}$$