Q1 (1) $f_1(t) = \lambda_1 e^{-\lambda_1 t}$

$S_1(t) = e^{-\lambda_1 t}$

$f_2(t) = \lambda_2 e^{-\lambda_2 t}$

$S_2(t) = e^{-\lambda_2 t}$

$$L(\lambda_1 \lambda_2 | t) = \prod_{i \in D} \lambda_1 e^{-\lambda_1 t_i} \prod_{j \in M} e^{-\lambda_1 t_j} \prod_{k \in B} \lambda_2 e^{-\lambda_2 t_k} \prod_{p \in A} e^{-\lambda_2 t_p}$$

D is the set of observation with exact time of Negative IM : $N_1 = |D|$
M is the set of Right censored observation of negative IM. : $N_2 = |M|$
B is the set of exact time of postive IM. : $N_3 = |B|$
A is the set of Right censored of positive IM. : $N_4 = |A|$

(2) $$\log L = N_1 \log\lambda_1 + \sum_{i=1}^{N_1} -\lambda_1 t_i + \sum_{j=1}^{N_2} -\lambda_1 t_j + N_3 \log\lambda_2 + \sum_{k=1}^{N_3} -\lambda_2 t_k + \sum_{p=1}^{N_4} -\lambda_2 t_p.$$

$$\frac{d \log L}{d\lambda_1} = \frac{1}{\lambda_1} \cdot N_1 - \sum t_i - \sum t_j = 0$$

$$\Rightarrow \lambda_1 = \frac{N_1}{\sum t_i + \sum t_j}$$

$\Rightarrow \lambda_1 = 0.005$

For the same reason $\Rightarrow \lambda_2 = \frac{N_3}{\sum t_k + \sum t_p}$

$\lambda_2 = 0.0056.$

(3) $S_1(36) = e^{-\lambda_1 \cdot 36} \approx 0.8594$

$S_2(36) = e^{-\lambda_2 \cdot 36} \approx 0.6298$

Q2 (1) Negative : 16

Positive : 8

(2) $\hat{S}_{Neg}(t) = \prod_{t_i \le 36} [1 - \frac{d_i}{Y_i}] = 0.889$

The 95% CI for Neg is [0.731, 0.957] ] Negative.

$\hat{S}_{Pos}(t) = 0.778$,

The 95% CI for Pos is [0.365, 0.939] ] positive

(3) The positive. median survival time is 73.
The negative median survival time is None.

This is Because more than one half of data are censored. and those censored data are all bigger than exact failure.

Scanned with CamScanner

(4) The NA estimator for $H(t)$ at 3 years for negative is.

$\hat{H}(t) = 0.1161$ , the standard error of it is 0.0581

The NA estimator for $H(t)$ at 3 years for positive is

$\hat{H}(t) = 0.2361$ , the standard error of it is 0.1672.

(5) ① The hazard function of negative samples is very closed with zero, ~~the~~ (show on other file)
This means that the instantaneous risk of this event at any time is very low.

② The hazard function of positive sample waves around time and significantly higher than zero.
This means that the instantaneous risk of positive IM at any time is very high.

③ it is not good to use exponential distribution. for positive IM Because of the wave
it is good to use exponential distribution for negative IM. Because it likes a line.

(6). For negative data: The risk of death for Negative IM is lower than positive IM.
The number of negative data is much more than positive samples. This may indicate that the Negative IM is truly at a low risk of death.

For positive data: There is a high risk of death when people are positive IM. Because of high risk of death, the number of positive samples are less than Negative samples.

Q3:

| Interval | Begin total | Adj ($Y_i$) | Death ($d_j$) | Loss ($m_j$) |
|---|---|---|---|---|
| (0-4] | 10 | 10 | 2 | 0 |
| (4-8] | 8 | 8-1.5 = 6.5 | 1 | 3 |
| (8-12] | 4 | 4 | 1 | 0 |
| (12-16] | 3 | 2.5 | 2 | 1 |

$\boxed{K\text{-}M}$

(1) $S(0) = 1$

$S(4) = (1-\frac{2}{10}) = 0.8 \qquad Var(S(4)) = 0.8^2 \cdot \frac{2}{10(10-2)} = 0.016$

$S(8) = 0.8 \times (1-\frac{1}{6.5}) = 0.6769. \qquad Var(S(8)) = (0.6769)^2 \cdot [\frac{2}{10\times 8} + \frac{1}{6.5(6.5-1)}] = 0.024$

$S(12) = 0.6769(1-\frac{1}{4}) = 0.5076 \qquad Var(S(12)) = (0.5076)^2 \cdot [\frac{2}{80} + \frac{1}{6.5\times 5.5} + \frac{1}{4(4-1)}] = 0.035$

$S(16) = 0.5076(1-\frac{2}{2.5}) = 0.10152 \qquad Var(S(16)) = (0.10152)^2 [\frac{2}{80} + \frac{1}{6.5\times 5.5} + \frac{1}{12} + \frac{2}{2.5\times 0.5}] = 0.0179.$

(2) $H(t) = -\ln[S(t)]$

$Var(H(t)) = Var(-\ln[S(t)]) = Var(\ln[S(t)]) \approx \sum_{t_i \le t} \frac{d_i}{Y_i(Y_i-d_i)}$

$H(0) = 0$

$H(4) = -\ln(\hat{S}(4)) = 0.2231 \qquad Var(H(4)) = 0.025$

$H(8) = -\ln(\hat{S}(8)) = 0.39 \qquad Var(H(8)) = 0.0529$

$H(12) = 0.5516 \qquad Var(H(12)) = 0.136.$

$H(16) = 2.287 \qquad Var(H(16)) = 1.736.$

$\boxed{N\text{-}A}$

(3) $H(0) = 0$

$H(4) = \frac{2}{10} = 0.25 \quad Var(H(4)) = \frac{2}{10^2} = 0.02.$

$H(8) = \frac{2}{10} + \frac{1}{6.5} = 0.2538 \quad Var(H(8)) = \frac{2}{10^2} + \frac{1}{6.5^2} = 0.04367$

$H(12) = 0.2538 + \frac{1}{4} = 0.5038 \quad Var(H(12)) = 0.04367 + \frac{1}{4^2} = 0.10617$

$H(16) = 0.5038 + \frac{2}{2.5} = 1.3038 \quad Var(H(16)) = 0.10617 + \frac{2}{2.5^2} = 0.42617.$

(4) $S(t) = e^{-H(t)} = g(Y_n)$ where $Y_n = H(t)$, $g(x) = e^{-x}$. By $\delta$-method, we have:

$Var(S(t)) = Var(e^{-H(t)}) \approx [e^{-H(t)}(-1)]^2 \cdot Var(H(t)) = [e^{-H(t)}]^2 \cdot Var(H(t))$

$S(0) = 1.$

$S(4) = 0.7788 \quad Var(S(4)) = (0.7788)^2 \cdot 0.02 = 0.0121$

$S(8) = 0.7758 \quad Var(S(8)) = (0.7758)^2 \cdot 0.04367 = 0.02628.$

$S(12) = 0.6042 \quad Var(S(12)) = (0.6042)^2 \cdot 0.10617 = 0.0387$

$S(16) = 0.2715 \quad Var(S(16)) = (0.2715)^2 \cdot 0.42617 = 0.0314$