NTNU

Norwegian University of
Science and Technology

Department of Mathematical Sciences

Examination paper for **TMA4275 Lifetime Analysis**

Academic contact during examination: Håkon Tjelmeland

Phone:

Examination date: June 2nd 2021 – solution sketch

Examination time (from–to): 09.00-13.00

Permitted examination support material: A

Other information:

Language: English

Number of pages: 10

Number of pages enclosed: 0

Checked by:

Informasjon om trykking av eksamensoppgave	
Originalen er:	
1-sidig ☐	**2-sidig** ☒
sort/hvit ☒	**farger** ☐
skal ha flervalgskjema ☐	

Date Signature

Note: Some of the exam problems were given in slightly different variants. This solution sketch gives only the solutions to the version A of each of the problems. The solutions of the other problem variants are similar to what is given here.

Problem 1

The density function is given from the survival function by

$$\begin{aligned} f(t) = -S'(t) &= -\frac{d}{dt}\left[\frac{1+\theta+2\theta t}{1+\theta}e^{-2\theta t}\right] \\ &= -\frac{1}{1+\theta}\left[2\theta e^{-2\theta t} + (1+\theta+2\theta t)e^{-2\theta t}\cdot(-2\theta)\right] \\ &= -\frac{e^{-2\theta t}}{1+\theta}\left[2\theta - 2\theta - 2\theta^2 - 4\theta^2 t\right] \\ &= \frac{e^{-2\theta t}}{1+\theta}\left(2\theta^2 + 4\theta^2 t\right) \\ &= \frac{2\theta^2(1+2t)}{1+\theta}e^{-2\theta t} \end{aligned}$$

So we have

$$\underline{\underline{f(t) = \frac{2\theta^2(1+2t)}{1+\theta}e^{-2\theta t} \text{ for } t \geq 0.}}$$

The hazard rate is given by

$$\begin{aligned} \alpha(t) = \frac{f(t)}{S(t)} &= \frac{\frac{2\theta^2(1+2t)}{1+\theta}e^{-2\theta t}}{\frac{1+\theta+2\theta t}{1+\theta}e^{-2\theta t}} \\ &= \frac{2\theta^2(1+2t)}{1+\theta+2\theta t}. \end{aligned}$$

So we have

$$\underline{\underline{\alpha(t) = \frac{2\theta^2(1+2t)}{1+\theta+2\theta t} \text{ for } t \geq 0.}}$$

Problem 2

- $P(dN(t) > 1) = 0$ is correct because a counting process is defined only to have jumps of size 1.

- The statement "$N(t)$ is Poisson distributed with mean value $\Lambda(t) = \int_0^t \lambda(s)ds$" is incorrect because $N(t)$ does not need to be Poisson distributed. If $N(t)$ counts the number of events in a Poisson process the statement would be true, but the statement is not generally correct.
- $N(t)$ is a sub-martingale. Since the $N(t)$ is an increasing function one must also have that $\mathrm{E}[N(t)|\mathcal{F}_s] \geq N(s)$.
- $2N(t) + 5$ is not a counting process because a counting process is required to start at zero at time zero.
- For $0 < t_1 < t_2 < t_3$ the increments $N(t_2) - N(t_1)$ and $N(t_3) - N(t_2)$ do not need to be independent. If $N(t)$ counts the number of events in a Poisson process it would be true, but the statement is not generally true for counting processes.
- It is not correct that $\lambda(t) \in [0, 1]$. The $\lambda(t)$ can not be negative, but it may be larger than one.

Problem 3

The Kaplan-Meyer estimates and corresponding confidence intervals may be found and plotted by the following R commands:

```
library(survival)
TPlacebo = c(1,1,2,2,3,4,4,5,5,8,8,8,8,11,11,12,12,15,17,22,23)
CensPlacebo = 1 - c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)


T6MP = c(6,6,6,6,7,9,10,10,11,13,16,17,19,20,22,23,25,32,32,34,35)
Cens6MP = 1 - c(0,0,0,1,0,1,0,1,1,0,0,1,1,1,0,0,1,1,1,1,1)


KMPlacebo = survfit(Surv(TPlacebo,CensPlacebo)~1,conf.type="log-log")
pdf("KMPlacebo.pdf")
plot(KMPlacebo,col="red")
graphics.off()


KM6MP = survfit(Surv(T6MP,Cens6MP)~1,conf.type="log-log")
pdf("KM6MP.pdf")
```

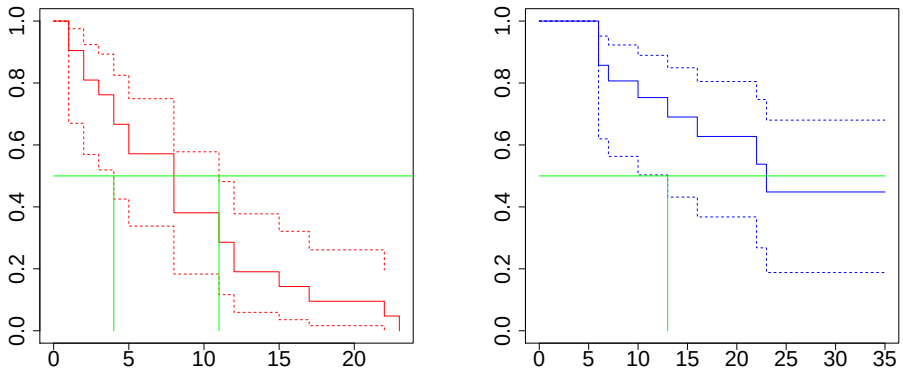

Figure 1: Kaplan-Meyer estimates with confidence intervals for the placebo (left) and the 6MP (right) groups.

plot(KM6MP,col="blue",new=FALSE)
graphics.off()

The resulting plots are shown in Figure 1. Confidence intervals for the median survival times can be found by drawing a horisontal line at survival probability equal to 0.5 and reading off at which times t this line crosses the confidence interval curves. The process is illustrated by green lines in the two plots in Figure 1 (code for producing these green lines for the placebo and 6Mp groups are

lines(c(0,25),c(0.5,0.5),col="green")
lines(c(4,4),c(0,0.5),col="green")
lines(c(11,11),c(0,0.5),col="green")

and

lines(c(0,35),c(0.5,0.5),col="green")
lines(c(13,13),c(0,0.5),col="green")

respectively). Note that for the 6MP group the upper limit in the confidence interval is equal to infinity. The numerical values for the confidence intervals can either be read off from the plots or they can be found by inspecting the R variables

KMPlacebo and KM6MP produced in the R code above, which can be done by the R commands

KMPlacenbo
KM6MP

The confidence interval for the Placebo group becomes $\underline{\underline{[4, 11]}}$ and for the 6MP group it becomes $\underline{\underline{[13, \infty)}}$.

Problem 4

The estimated relative risk function is

$$\underline{\underline{r(t, x) = \exp\{-0.51 \cdot \text{sex} + 0.015 \cdot \text{age} - 0.0022 \cdot \text{wt.loss} - 0.013 \cdot \text{ph.karno}\}}}$$

From the R output we see that $\underline{\underline{\text{the two covariates sex and ph.karno are significant at the 5\% level}}}$ since these two have values $Pr(>|z|)$ that are smaller than 0.05.

We are asked to find a 95% confidence interval for the ratio

$$\frac{\alpha_{\text{male}}(t|x)}{\alpha_{\text{female}}(t|x)} = \frac{\alpha_0(t)\exp\{\beta_1 \cdot 1 + \beta_2 \cdot x_2 + \beta_3 x_3 + \beta_4 x_4\}}{\alpha_0(t)\exp\{\beta_1 \cdot 2 + \beta_2 \cdot x_2 + \beta_3 x_3 + \beta_4 x_4\}} = \exp\{-\beta_1\}.$$

From the R output we find that a 95% confidence interval for β_1 is

$$(-0.513955 - z_{0.025} \cdot 0.174410, -0.513955 + z_{0.025} \cdot 0.174410) = (-0.8557986, -0.1721114),$$

where we used that $z_{0.025} = 1.96$. The 95% confidence interval for $e^{-\beta_1}$ then becomes

$$(\exp\{-(-0.1721114)\}, \exp\{-(-0.857986)\}) = \underline{\underline{(1.1878, 2.3527)}}$$

To see what covariates that have the largest effect on the survival probability we need to take into account over which interval the various covariates are varying. These intervals are given in the beginning of the problem text and we get:

- sex: $|-0.513955 \cdot (2-1))| = 0.513955$
- age: $|0.015140 \cdot (82-39)| = 0.65102$
- wt.loss: $|-0.002246 \cdot (68-(-24))| = 0.206632$

- ph.karno: $|-0.012871 \cdot (100-50)| = 0.64355$

So we see that in the estimated model age is the covariate that has the largest effect on the hazard rate, and thereby also the largest effect on the survival probability.

What to do next in the analysis of the data set? Since two of the covariates are not significant is would be reasonable to start trying a model with either age or wt.loss removed, or both age and wt.loss removed. Thereafter is would be natural to check how well the estimated model fit the data by looking at for example the martingale residuals and by comparing the fitted model with the results for a stratified model, where for example the two sexes are allowed to have different baseline hazard rates.

Problem 5

- $8M_n$ is a mean zero martingale. It is easy to show that multiplying with a constant preserves the martingale property, and clearly also $8M_0 = 0$ since $M_0 = 0$.
- $2M_n + 4$ is not a mean zero martingale since $2M_0 + 4 = 4 \neq 0$. So even if the martingale property is preserved for $2M_n + 4$ it is not starting at zero.
- $M_n^2 + M_n$ is not (necessarily) a martingale. The easiest way to show this is to select a particular zero mean martingale and show that for that process the martingale property does not hold for $M_n^2 + M_n$.
- $(H \bullet M)_n$ is a mean zero martingale. This is in fact an important result in our textbook.
- $(H^2 \bullet M)_n$ is a mean zero martingale because when H_n is a predictable process also H_n^2 becomes a predictable process, and thereby the previous result gives that also $(H^2 \bullet M)_n$ is a zero mean martingale.
- $\sum_{s=1}^{n}(2s+1)H_s(M_s - M_{s-1})$ is also a zero mean martingale. This is because when H_n is a predictable process, then also $\widetilde{H}_n = (2n+1)H_n$ is a predictable process, and as we then have that

  $$\sum_{s=1}^{n}(2s+1)H_s(M_s - M_{s-1}) = (\widetilde{H} \bullet M)_n$$

  the result follows.

Problem 6

To show that $X(t)$ is a sub-martingale with respect to $\{\mathcal{F}_t\}$ we need to show that for any $0 \leq s < t$ we have

$$\mathrm{E}[X(t)|\mathcal{F}_s] \geq X(s).$$

Inserting how $X(t)$ is defined in $\mathrm{E}[X(t)|\mathcal{F}_s]$ we get

$$\begin{aligned}
\mathrm{E}[X(t)|\mathcal{F}_s] &= \mathrm{E}\left[\sum_{i=1}^{N(t)} Z_i \middle| \mathcal{F}_s\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{N(s)} Z_i + \sum_{i=N(s)+1}^{N(t)} \middle| \mathcal{F}_s\right] \\
&= \sum_{i=1}^{N(s)} Z_i + \mathrm{E}\left[\sum_{i=N(s)+1}^{N(t)} Z_i \middle| \mathcal{F}_s\right] \\
&= X(s) + \mathrm{E}\left[\sum_{i=N(s)+1}^{N(t)} Z_i \middle| \mathcal{F}_s\right]
\end{aligned}$$

Now focusing on the last term in this extression and using the law of double expectation by conditioning on $N(t)$ we get

$$\begin{aligned}
\mathrm{E}\left[\sum_{i=N(s)+1}^{N(t)} Z_i \middle| \mathcal{F}_s\right] &= \mathrm{E}\left[\mathrm{E}\left[\sum_{i=N(s)+1}^{N(t)} Z_i \middle| \mathcal{F}_s, N(t)\right] \middle| \mathcal{F}_s\right] \\
&= \mathrm{E}\left[\sum_{i=N(s)+1}^{N(t)} \mathrm{E}[Z_i|\,\mathcal{F}_s, N(t)] \middle| \mathcal{F}_s\right] \\
&= \mathrm{E}\left[\sum_{i=N(s)+1}^{N(t)} \mu \middle| \mathcal{F}_s\right] \\
&= \mathrm{E}[\mu(N(t) - N(s))|\,\mathcal{F}_s] \\
&= \mu \mathrm{E}[N(t) - N(s)|\,\mathcal{F}_s]\,.
\end{aligned}$$

In the problem text it is given that $\mu > 0$, and since $N(t)$ is a counting process we have that $N(t) - N(s) \geq 0$ and thereby also $\mathrm{E}[N(t) - N(s)|\mathcal{F}_s] \geq 0$. Thereby we have that

$$\mathrm{E}\left[\sum_{i=N(s)+1}^{N(t)} Z_i \middle| \mathcal{F}_s\right] \geq 0,$$

and thereby also

$$\mathrm{E}[X(t)|\mathcal{F}_s] \geq X(s)$$

as we should show.

To find the compensator $X^\star(t)$ we can use that

$$dX^\star(t) = \mathrm{E}[dX(t)|\mathcal{F}_{t-}]$$

and that the expression we have for $X(t)$ gives that

$$dX(t) = X(t) - X(t-) = I(dN(t) = 1)Z_{N(t)}.$$

Combining these two expressions we get

$$dX^\star(t) = \mathrm{E}[I(dN(t) = 1)Z_{N(t)}|\mathcal{F}_{t-}].$$

Using the law of double expectation, conditioning on the value of $dN(t)$, we get

$$\begin{aligned}
dX^\star(t) &= \mathrm{E}[I(dN(t) = 1)Z_{N(t)}|\mathcal{F}_{t-}, dN(t) = 0] \cdot P(dN(t) = 0|\mathcal{F}_{t-}) \\
&\quad + \mathrm{E}[I(dN(t) = 1)Z_{N(t)}|\mathcal{F}_{t-}, dN(t) = 1] \cdot P(dN(t) = 1|\mathcal{F}_{t-}) \\
&= \mathrm{E}[0 \cdot Z_{N(t)}|\mathcal{F}_{t-}, dN(t) = 0] \cdot P(dN(t) = 0|\mathcal{F}_{t-}) \\
&\quad + \mathrm{E}[1 \cdot Z_{N(t)}|\mathcal{F}_{t-}, dN(t) = 1] \cdot P(dN(t) = 1|\mathcal{F}_{t-}) \\
&= 0 + \mathrm{E}[Z_{N(t)}] \cdot \lambda(t)dt = \mu\lambda(t)dt,
\end{aligned}$$

where we in the last line used that when it is given that there is a new event at time t, the value of $Z_{N(t)}$ is independent of everything that has happened before time t, and that the intensity process $\lambda(t)$ can be expressed as

$$\lambda(t)dt = P(dN(t) = 1|\mathcal{F}_{t-}) = \mathrm{E}[dN(t) = 1|\mathcal{F}_{t-}].$$

Thereby we have that the compensator becomes

$$X^\star(t) = \int_0^t dX^\star(s)ds = \mu \int_0^t \lambda(s)ds = \mu\Lambda(t).$$

The martingale $M(t)$ in the Doob-Meyer decomposition then becomes

$$M(t) = X(t) - X^\star(t) = \sum_{i=1}^{N(t)} Z_i - \mu\Lambda(t).$$

Problem 7

a)

The intensity process for individual number i becomes

$$\lambda_i(t;\theta) = Y_i(t)\alpha_i(t) = I(\widetilde{T}_i \geq t)\alpha_i(t) = I(\widetilde{T}_i \geq t)\nu \exp\{\beta_1 x_{i1} + \beta_2 x_{i2}\}.$$

Thereby the aggregated intensity process becomes

$$\begin{aligned}\lambda_\bullet(t;\theta) &= \sum_{i=1}^n \lambda_i(t;\theta) \\ &= \sum_{i=1}^n I(\widetilde{T}_i \geq t)\nu \exp\{\beta_1 x_{i1} + \beta_2 x_{i2}\}. \\ &= \nu \sum_{i=1}^n I(\widetilde{T}_i \geq t) \exp\{\beta_1 x_{i1} + \beta_2 x_{i2}\}.\end{aligned}$$

Using that we necessarily must have $I(\widetilde{T}_i \geq t) = 1$ whenever $\Delta N_i(t) = 1$ we thereby get from the general formula for the likelihood function of a counting process that

$$\begin{aligned}L(\theta) &= \prod_{i=1}^n \left(\nu \exp\{\beta_1 x_{i1} + \beta_2 x_{i2}\}\right)^{D_i} \exp\left\{-\int_0^\tau \nu \sum_{i=1}^n I(\widetilde{T}_i \geq t) \exp\{\beta_1 x_{i1} + \beta_2 x_{i2}\} dt\right\} \\ &= \nu^{D_\bullet} \exp\left\{\sum_{i=1}^n D_i(\beta_1 x_{i1} + \beta_2 x_{i2})\right\} \exp\left\{-\nu \sum_{i=1}^n \exp\{\beta_1 x_{i1} + \beta_2 x_{i2}\} \int_0^\tau I(\widetilde{T}_i \geq t) dt\right\} \\ &= \nu^{D_\bullet} \exp\left\{\beta_1 \sum_{i=1}^n D_i x_{i1} + \beta_2 \sum_{i=1}^n D_i x_{i2}\right\} \exp\left\{-\nu \sum_{i=1}^n e^{\beta_1 x_{i1} + \beta_2 x_{i2}} \widetilde{T}_i\right\} \\ &= \nu^{D_\bullet} \exp\left\{\beta_1 \sum_{i:x_{i1}=1} D_i + \beta_2 \sum_{i=1}^n D_i x_{i2} - \nu \left(\sum_{i:x_{i1}=0} e^{\beta_2 x_{i2}} \widetilde{T}_i + \sum_{i:x_{i1}=1} e^{\beta_1 + \beta_2 x_{i2}} \widetilde{T}_i\right)\right\} \\ &= \nu^{D_\bullet} \exp\left\{\beta_1 D_\bullet^{(1)} + \beta_2 \sum_{i=1}^n D_i x_{i2} - \nu \left(\sum_{i:x_{i1}=0} e^{\beta_2 x_{i2}} \widetilde{T}_i + e^{\beta_1} \sum_{i:x_{i1}=1} e^{\beta_2 x_{i2}} \widetilde{T}_i\right)\right\},\end{aligned}$$

where we in the second last row used that $X_{i1} \in \{0, 1\}$. The log likelihood function thereby becomes

$$\ell(\theta) = D_\bullet \ln \nu + \beta_1 D_\bullet^{(1)} + \beta_2 \sum_{i=1}^n D_i x_{i2} - \nu \left(\sum_{i:x_{i1}=0} e^{\beta_2 x_{i2}} \widetilde{T}_i + e^{\beta_1} \sum_{i:x_{i1}=1} e^{\beta_2 x_{i2}} \widetilde{T}_i\right)$$

which is what we should show.

b) The partial derivatives of $\ell(\theta)$ with respect to ν and β_1, repectively, become

$$\frac{\partial \ell}{\partial \nu} = D_\bullet \cdot \frac{1}{\nu} - \left(\sum_{i:x_{i1}=0} e^{\beta_2 x_{i2}} \widetilde{T}_i + e^{\beta_1} \sum_{i:x_{i1}=1} e^{\beta_2 x_{i2}} \widetilde{T}_i \right),$$
$$\frac{\partial \ell}{\partial \beta_1} = D_\bullet^{(1)} - \nu e^{\beta_1} \sum_{i:x_{i1}=1} e^{\beta_2 x_{i2}}.$$

Defining the notations

$$S^{(0)}(\beta_2) = \sum_{i:x_{i1}=0} e^{\beta_2 x_{i2}} \widetilde{T}_i,$$
$$S^{(1)}(\beta_2) = \sum_{i:x_{i1}=1} e^{\beta_2 x_{i2}} \widetilde{T}_i,$$

and solving $\frac{\partial \ell}{\partial \nu} = 0$ with respect to ν we get

$$\nu = \frac{D_\bullet}{S^{(0)}(\beta_2) + e^{\beta_1} S^{(1)}(\beta_2)}.$$

Then setting $\frac{\partial \ell}{\partial \beta_1} = 0$ and inserting the expression we just found for ν we get

$$\begin{aligned}
D_\bullet^{(1)} &= \frac{D_\bullet e^{\beta_1} S^{(1)}(\beta_2)}{S^{(0)}(\beta_2) + e^{\beta_1} S^{(1)}(\beta_2)} \\
D_\bullet^{(1)} S^{(0)}(\beta_2) + e^{\beta_1} D_\bullet^{(1)} S^{(1)}(\beta_2) &= e^{\beta_1} D_\bullet S^{(1)}(\beta_2) \\
e^{\beta_1} &= \frac{D_\bullet^{(1)} S^{(0)}(\beta_2)}{D_\bullet S^{(1)}(\beta_2) - D_\bullet^{(1)} S^{(1)}(\beta_2)} \\
\beta_1 &= \ln \left[\frac{D_\bullet^{(1)} S^{(0)}(\beta_2)}{D_\bullet S^{(1)}(\beta_2) - D_\bullet^{(1)} S^{(1)}(\beta_2)} \right].
\end{aligned}$$

We have thereby explicit expressions for the MLEs for ν and β_1 as a function of the MLE for β_2,

$$\widehat{\beta}_1 = \ln \left[\frac{D_\bullet^{(1)} S^{(0)}(\widehat{\beta}_2)}{D_\bullet S^{(1)}(\widehat{\beta}_2) - D_\bullet^{(1)} S^{(1)}(\widehat{\beta}_2)} \right]$$
$$\widehat{\nu} = \frac{D_\bullet}{S^{(0)}(\widehat{\beta}_2) + e^{\widehat{\beta}_1} S^{(1)}(\widehat{\beta}_2)},$$

and we get the profile likelihood for β_2 by inserting the expressions we have found for ν and β_1 as a function of β_2 in the formula we have for $\ell(\theta)$.

c)

We know that we approximately have that $\widehat{\theta} \sim N(0, \mathbb{I}(\widehat{\theta}))$. Under H_0 we thereby get that $\widehat{\beta_1}$ is approximately normal with zero mean and variance equal to

$$-\left.\frac{\partial^2 \ell}{\beta_1^2}\right|_{\theta=\widehat{\theta}} = \widehat{\nu} e^{\widehat{\beta}_1} S^{(1)}(\widehat{\beta}_2).$$

As a test statistic we can thereby use

$$Z = \frac{\widehat{\beta}_1}{\sqrt{\widehat{\nu} e^{\widehat{\beta}_1} S^{(1)}(\widehat{\beta}_2)}}$$

which is approximately standard normal when H_0 is true. We therefore should reject H_0 at signifiance level α if $|Z| > z_{\frac{\alpha}{2}}$.