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: May 21st 2024 – solution sketch

Examination time (from–to): 15.00-19.00

Permitted examination support material: C:

Tabeller og formler i statistikk, Akademika
A yellow sheet of paper (A5 with a stamp) with personal handwritten formulas and notes
A specific basic calculator

Other information:

Language: English

Number of pages: 9

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

Problem 1

We first use the given hazard rate $\alpha(t)$ to find the integrated hazard rate $A(t)$ and the survival function $S(t)$,

$$\begin{aligned} A(t) &= \int_0^t \alpha(u)du = \int_0^t 2u^{\frac{1}{2}}du = \left[\frac{4}{3}u^{\frac{3}{2}}\right]_0^t = \frac{4}{3}t^{\frac{3}{2}}, \\ S(t) &= P(T > t) = e^{-A(t)} = e^{-\frac{4}{3}t^{\frac{3}{2}}}. \end{aligned}$$

Using the formula for $S(t)$ we can then find the two probabilities,

$$\begin{aligned} P(T > 1.5) &= S(1.5) = e^{-\frac{4}{3}\cdot 1.5^{\frac{3}{2}}} = \underline{\underline{0.086338}}, \\ P(T \leq 2 | T > 1.5) &= \frac{P(T \leq 2 \cap T > 1.5)}{P(T > 1.5)} = \frac{P(1.5 < T \leq 2)}{P(T > 1.5)} = \frac{P(T \leq 2) - P(T \leq 1.5)}{P(T > 1.5)} \\ &= \frac{(1 - P(T > 2)) - (1 - P(T > 1.5))}{P(T > 1.5)} = \frac{P(T > 1.5) - P(T > 2)}{P(T > 1.5)} \\ &= \frac{S(1.5) - S(2)}{S(1.5)} = \frac{\exp\left\{-\frac{4}{3}\cdot 1.5^{\frac{3}{2}}\right\} - \exp\left\{-\frac{4}{3}\cdot 2^{\frac{3}{2}}\right\}}{\exp\left\{-\frac{4}{3}\cdot 1.5^{\frac{3}{2}}\right\}} = \underline{\underline{0.733331}}. \end{aligned}$$

Letting m denote the median in the distribution of T, the value of m is given by

$$\begin{aligned} S(m) &= \frac{1}{2} \\ e^{-\frac{4}{3}m^{\frac{3}{2}}} &= \frac{1}{2} \\ -\frac{4}{3}m^{\frac{3}{2}} &= -\ln(2) \\ m^{\frac{3}{2}} &= \frac{3}{4}\ln(2) \\ m &= \left(\frac{3}{4}\ln(2)\right)^{\frac{2}{3}} = \underline{\underline{0.646534}}. \end{aligned}$$

Problem 2 The estimated risk function is

$$\underline{\underline{r(x, \hat{\beta}) = \exp\{0.04681 \cdot x_1 - 0.29279 \cdot x_2 + 0.15813 \cdot x_3\}}}.$$

The hazard rate of an individual with $x = [x_1, -2, x_3]^T$ is

$$\alpha(t|x) = \alpha_0(t) \cdot \exp\{\beta_1 x_1 + \beta_2 \cdot (-2) + \beta_3 x_3\},$$

and the hazard rate of an individual with $\widetilde{x} = [x_1, 2, x_3]^T$ is

$$\alpha(t|x) = \alpha_0(t) \cdot \exp\{\beta_1 x_1 + \beta_2 \cdot 2 + \beta_3 x_3\}.$$

When the values of x_1 and for x_3 are equal for the two individuals, the ratio between the two hazard rates becomes

$$\frac{\alpha(t|x)}{\alpha(t|\widetilde{x})} = \frac{e^{-2\beta_2}}{e^{2\beta}} = e^{-4\beta_2}.$$

From the R output we see that a 95% confidence interval for β_2 is

$$[-0.29279 - 1.96 \cdot 0.09810, -0.29279 + 1.96 \cdot 0.09810] = [-0.485066, -0.100514].$$

Noting that $e^{-4\beta_2}$ is a strictly decreasing function of β_2, the corresponding 95% confidence interval for $e^{-4\beta_2}$ becomes

$$\left[e^{-4\cdot(-0.100514)}, e^{-4\cdot(-0.485066)}\right] = \underline{\underline{[1.494895, 6.960588]}}.$$

What to do next in the analysis of the data set? We see that the covariate x_1 is clearly not significant, whereas the covariate x_3 is in the border of being significant. It would therefore be natural to try a model where the covariate x_1 is removed. Then there are two possibilities, either the covariate x_3 is significant in such a model, or x_3 is not significant. If x_3 is not significant in such a reduced a model it would be natural to remove also x_3 from the model. Thereafter, having settled on a model with either both x_2 and x_3 as covariates or a model with only x_2, it would be natural to check how well the estimated model fit the data by looking at for example the martingale residuals.

Problem 3

a) We let k_A and k_B denote the number of observed failures for units of type A and type B, respectively. The observed failure times for units of type A we denote by $T_1^A, \dots, T_{k_A}^A$, and the observed failure times for units of type B we correspondingly denote by $T_1^B, \dots, T_{k_B}^B$.

The likelihood function is then given by

$$
\begin{aligned}
L(\alpha,\beta,\gamma) &= \left[\prod_{j=1}^{k_A} \lambda_A(T_j^A)\right] \cdot \left[\prod_{j=1}^{k_B} \lambda_B(T_j^B)\right] \cdot \exp\left\{-\int_0^\tau (n\lambda_A(t) + m\lambda_B(t))dt\right\} \\
&= \left[\prod_{j=1}^{k_A} \alpha(T_j^A)^{\gamma-1}\right] \cdot \left[\prod_{j=1}^{k_B} \beta(T_j^B)^{\gamma-1}\right] \cdot \exp\left\{-\int_0^\tau n\alpha t^{\gamma-1} + m\beta t^{\gamma-1} dt\right\} \\
&= \left[\prod_{j=1}^{k_A} \alpha(T_j^A)^{\gamma-1}\right] \cdot \left[\prod_{j=1}^{k_B} \beta(T_j^B)^{\gamma-1}\right] \cdot \exp\left\{-\frac{n\alpha}{\gamma}\tau^\gamma - \frac{m\beta}{\gamma}\tau^\gamma\right\}
\end{aligned}
$$

The log-likelihood function becomes

$$
\begin{aligned}
\ell(\alpha,\beta,\gamma) &= \ln L(\alpha,\beta,\gamma) \\
&= \sum_{j=1}^{k_A}\left[\ln\alpha + (\gamma-1)\ln T_j^A\right] + \sum_{j=1}^{k_B}\left[\ln\beta + (\gamma-1)\ln T_j^B\right] - \frac{n\alpha}{\gamma}\tau^\gamma - \frac{m\beta}{\gamma}\tau^\gamma \\
&= k_A\ln\alpha + (\gamma-1)\sum_{j=1}^{k_A}\ln T_j^A + k_B\ln\beta + (\gamma-1)\sum_{j=1}^{k_B}\ln T_j^B - \frac{n\alpha}{\gamma}\tau^\gamma - \frac{m\beta}{\gamma}\tau^\gamma \\
&= -\left[\sum_{j=1}^{k_A}\ln T_j^A + \sum_{j=1}^{k_B}\ln T_j^B\right] + k_A\ln\alpha + k_B\ln\beta + \gamma\left[\sum_{j=1}^{k_A}\ln T_j^A + \ln_{j=1}^{k_B}\ln T_j^B\right] \\
&- \frac{n\alpha}{\gamma}\tau^\gamma - \frac{m\beta}{\gamma}\tau^\gamma \\
&= s_0 + s_1\ln\alpha + s_2\ln\beta + s_3\gamma - \frac{n\alpha}{\gamma}\tau^\gamma - \frac{m\beta}{\gamma}\tau^\gamma,
\end{aligned}
$$

where

$$
\begin{aligned}
s_0 &= -\left[\sum_{j=1}^{k_A}\ln T_j^A + \sum_{j=1}^{k_B}\ln T_j^B\right], \\
s_1 &= k_A, \\
s_2 &= k_B, \\
s_3 &= \left[\sum_{j=1}^{k_A}\ln T_j^A + \sum_{j=1}^{k_B}\ln T_j^B\right].
\end{aligned}
$$

b) The partial derivatives of the log likelihood function becomes

$$\begin{aligned}
\frac{\partial \ell}{\partial \alpha} &= \frac{s_1}{\alpha} - \frac{n}{\gamma}\tau^{\gamma}, \\
\frac{\partial \ell}{\partial \beta} &= \frac{s_2}{\beta} - \frac{m}{\gamma}\tau^{\gamma}, \\
\frac{\partial \ell}{\partial \gamma} &= s_3 - \left[-\frac{n\alpha}{\gamma^2}\tau^{\gamma} + \frac{n\alpha}{\gamma}\tau^{\gamma}\ln\tau\right] - \left[-\frac{m\beta}{\gamma^2}\tau^{\gamma} + \frac{m\beta}{\gamma}\tau^{\gamma}\ln\tau\right].
\end{aligned}$$

We see that we easily can solve $\frac{\partial \ell}{\partial \alpha} = 0$ and $\frac{\partial \ell}{\partial \beta} = 0$ with respect to α and β, respectively, and thereby express α and β in terms of γ. It is more unclear what we can do with the partial derivative with respect to γ. Starting with setting the partial derivatives with respect to α and β equal to zero, we get

$$\begin{aligned}
\frac{\partial \ell}{\partial \alpha} = 0 \Leftrightarrow \alpha = \frac{s_1\gamma}{n}\tau^{-\gamma}, \\
\frac{\partial \ell}{\partial \beta} = 0 \Leftrightarrow \beta = \frac{s_2\gamma}{m}\tau^{-\gamma}.
\end{aligned}$$

Inserting these expressions for α and β in the expression for the log likelihood function we get the profile likelihood for γ,

$$\begin{aligned}
\ell_p(\gamma) &= \ell\left(\frac{s_1\gamma}{n}\tau^{-\gamma}, \frac{s_2\gamma}{m}\tau^{-\gamma}, \gamma\right) \\
&= s_0 + s_1(\ln s_1 + \ln\gamma - \ln n - \gamma\ln\tau) + s_2(\ln s_2 + \ln\gamma - \ln m - \gamma\ln\tau) \\
&+ s_3\gamma - \frac{n}{\gamma}\cdot\frac{s_1\gamma}{n}\tau^{-\gamma}\cdot\tau^{\gamma} - \frac{m}{\gamma}\cdot\frac{s_2\gamma}{m}\tau^{-\gamma}\cdot\tau^{\gamma} \\
&= s_4 + s_5\gamma + s_6\ln\gamma,
\end{aligned}$$

where

$$\begin{aligned}
s_4 &= s_0 + s_1\ln s_1 + s_2\ln s_2 - s_1\ln n - s_2\ln m - s_1 - s_2, \\
s_5 &= s_3 - (s_1 + s_2)\ln\tau, \\
s_6 &= s_1 + s_2.
\end{aligned}$$

Now we can note that this profile likelihood can be optimised analytically. We have that

$$\ell_p'(\gamma) = s_5 + \frac{s_6}{\gamma},$$

which gives

$$\ell_p'(\gamma) = 0 \Leftrightarrow \gamma = -\frac{s_6}{s_5}.$$

The maximum likelihood estimators for α, β and γ thereby becomes

$$\begin{aligned}
\widehat{\gamma} &= -\frac{s_6}{s_5}, \\
\widehat{\alpha} &= \frac{s_1\widehat{\gamma}}{n}\tau^{-\widehat{\gamma}} = -\frac{s_1 \cdot s_6}{ns_5}\tau^{-\frac{s_6}{s_5}}, \\
\widehat{\beta} &= \frac{s_2\widehat{\gamma}}{m}\tau^{-\widehat{\gamma}} = -\frac{s_2 \cdot s_6}{ms_5}\tau^{-\frac{s_6}{s_5}}.
\end{aligned}$$

One should note that

$$\begin{aligned}
\widehat{\gamma} = -\frac{s_6}{s_5} &= \frac{1}{\frac{(s_1+s_2)\ln\tau - s_3}{s_1+s_2}} \\
&= \frac{1}{\ln\tau - \frac{\sum_{j=1}^{k_A}\ln T_j^A + \sum_{j=1}^{k_b}\ln T_j^B}{k_A+k_b}} \geq 0,
\end{aligned}$$

because $\ln\tau \geq \ln T_j^A$ and $\ln\tau \geq \ln T_j^B$ for all j. So we have $\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma} \geq 0$.

Problem 4

a) On incremental form the Doob-Meyer decomposition of $N(t)$ is

$$dN(s) = \lambda(s)ds + dM(s) = Y(s)\alpha(s)ds + dM(s).$$

Using this we get that

$$\begin{aligned}
\widehat{A}(t) &= \int_0^t \frac{J(s)}{Y(s)}dN(s) = \int_0^t \frac{J(s)}{Y(s)}(Y(s)\alpha(s)ds + dM(s)) \\
&= \int_0^t J(s)\alpha(s)ds + \int_0^t \frac{J(s)}{Y(s)}dM(s).
\end{aligned}$$

Combining this with the definition of $A_0^\star(t)$ we get that

$$\begin{aligned}
\widehat{A}(t) - A_0^\star(t) &= \int_0^t J(s)\alpha(s)ds + \int_0^t \frac{J(s)}{Y(s)}dM(s) - \int_0^t J(s)\alpha_0(s)ds \\
&= \int_0^t J(s)(\alpha(s) - \alpha_0(s))ds + \int_0^t \frac{J(s)}{Y(s)}dM(s),
\end{aligned}$$

as we should show. On incremental form this expression reads

$$d\widehat{A}(s) - dA_0^\star(s) = J(s)(\alpha(s) - \alpha_0(s))ds + \frac{J(s)}{Y(s)}dM(s),$$

which gives

$$\begin{aligned} Z(t_0) &= \int_0^{t_0} Y(s)\left(d\widehat{A}(s) - dA_0^\star(s)\right) \\ &= \int_0^{t_0} Y(s)\left(J(s)(\alpha(s)-\alpha_0(s))ds + \frac{J(s)}{Y(s)}dM(s)\right) \\ &= \underline{\underline{\int_0^{t_0} J(s)Y(s)(\alpha(s)-\alpha_0(s))ds + \int_0^{t_0} J(s)dM(s).}} \end{aligned}$$

When H_0 is true we have $\alpha(s) = \alpha_0(s)$ for all $s \in [0, t_0]$ so that the integrand in the first integral is identical to zero. Thereby, when H_0 is true we have

$$Z(t_0) = \int_0^{t_0} J(s)dM(s).$$

By definition $Y(s)$ is a predictable function with respect to the history $\mathcal{F}_t$. Thereby also the indicator function $J(s) = \mathbf{I}(Y(s) > 0$ becomes predictable with respect to the same history, which implies that $Z(t_0)$ is a stochastic integral with respect to a mean zero martingale. The $Z_(t_0)$ is thereby also a mean zero martingale when H_0 is true, which in particular implies that

$$E[Z(t_0)] = 0.$$

b) Assuming H_0 to be true we have from the above that

$$Z(t_0) = \int_0^{t_0} J(s)dM(s),$$

where $M(s)$ is a counting process martingale. This gives that

$$\langle Z\rangle(t_0) = \int_0^{t_0} (J(s))^2 d\langle M\rangle(s).$$

Since $J(s)$ is an indicator function we have that $(J(s))^2 = J(s)$, so that

$$\langle Z\rangle(t_0) = \int_0^{t_0} J(s)d\langle M\rangle(s).$$

Since $M(s)$ is a counting process martingale we have that

$$\langle M\rangle(t) = \int_0^t \lambda(s)ds = \int_0^t Y(s)\alpha(s)ds,$$

which in incremental form reads

$$d\langle M\rangle(s) = Y(s)\alpha(s)ds = Y(s)\alpha_0(s)ds,$$

where we have used that $\alpha(s) = \alpha_0(s)$ for all $s \in [0, t_0]$ when H_0 is true. Inserting this last expression in our expression for $Z(t_0)$ we get

$$\langle Z\rangle(t_0) = \int_0^{t_0} J(s)Y(s)\alpha_0(s)ds.$$

We have that $J(s)Y(s) = Y(s)$ since $J(s)$ equals zero only if $Y(s) = 0$, and is one otherwise. So we have

$$\underline{\underline{\langle Z\rangle(t_0) = \int_0^{t_0} Y(s)\alpha_0(s)ds.}}$$

As we know

$$\text{Var}[Z(t_0)] = E[\langle Z\rangle(t_0)]$$

this gives that

$$\text{Var}[Z(t_0)] = E\left[\int_0^{t_0} Y(s)\alpha_0(s)ds\right],$$

and thereby an unbiased estimator for $\text{Var}[Z(t_0)]$ is

$$\underline{\underline{\widehat{\text{Var}[Z(t_0)]} = \int_0^{t_0} Y(s)\alpha_0(s)ds.}}$$

Using martingale limit theorems one can show that $Z(t_0)$ is approximately normal. As test statistic one therefore use

$$\frac{Z(t_0)}{\sqrt{\int_0^{t_0} Y(s)\alpha_0(s)ds}},$$

which one can show is approximately standard normal. So one reject H_0 if the absolute value of this test statistic is sufficiently much different from zero.

Problem 5

a) The Nelson-Aalen estimator $\widehat{A}_{01}(t)$ is a step function starting at zero at time $t = 0$ and having steps whenever a transition from state 0 to state 1 is observed. In the time interval $t \in [0, 1]$ transitions from state 0 to state 1 is observed at the times

$$0.11(\text{unit } 4), 0.53(\text{unit } 1), 0.92(\text{unit } 4).$$

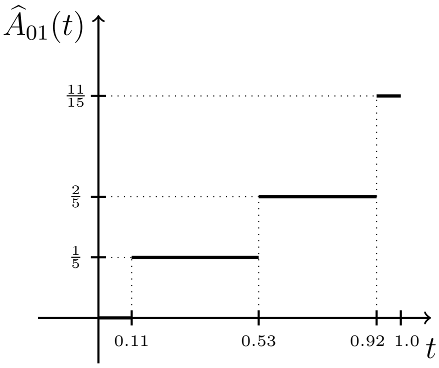

Figure 1: A plot of the Nelson-Aalen estimator $\widehat{A}_{01}(t)$ for $t \in [0, 1.0]$.

The $\widehat{A}_{01}(t)$ therefore have jumps for $t \in \{0.11, 0.53, 0.92\}$. To decide the height of the jumps we need to decide the number of units that are in state 0, $Y_0(t)$, just before each at these times. Just before time $t = 0.11$ all the units are still in state 0, so $Y_0(0.11) = 5$. Just before time $t = 0.53$ all the units are also in state 0 (since unit 4 has returned to state 0 before this time), so $Y_0(0.53) = 5$. Just before time $t = 0.92$ we see that unit 1 is in state 1, states 2 and 3 are still in state 0, unit 4 is in state 0 and unit 5 is in state 2, so $Y_0(0.92) = 3$. Thereby we have

$$\begin{aligned}
\widehat{A}_{01}(t) &= 0 \quad \text{for } t \in [0, 0.11), \\
\widehat{A}_{01}(t) &= \frac{1}{5} \quad \text{for } t \in [0.11, 0.53), \\
\widehat{A}_{01}(t) &= \frac{1}{5} + \frac{1}{5} = \frac{2}{5} \quad \text{for } t \in [0.53, 0.92), \\
\widehat{A}_{01}(t) &= \frac{1}{5} + \frac{1}{5} + \frac{1}{3} = \frac{11}{15} \quad t \in [0.92, 1.0].
\end{aligned}$$

A plot of $\widehat{A}_{01}(t)$ for $t \in [0, 1.0]$ is shown in Figure 1.

b) To find $\widehat{\mathbf{P}}(0.50, 0.75)$ we need first to find all event times in the time interval $t \in (0.50, 1.0]$. These event times are $t = 0.53$ and 0.57. Thereby we have

$$\widehat{\mathbf{P}}(0.50, 0.75) = (\mathbf{I} + \Delta\widehat{\mathbf{A}}(0.53))(\mathbf{I} + \Delta\widehat{\mathbf{A}}(0.57))$$

At time $t = 0.53$ we observe a transition from state 0 to state 1, and we have as previously discussed $Y_0(0.53) = 5$, so

$$\Delta\widehat{\mathbf{A}}(0.53) = \begin{bmatrix} -\frac{1}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

At time $t = 0.57$ we observe at transition from state 0 to state 2, and we have $Y_0(0.57) = 4$, so

$$\Delta\widehat{\mathbf{A}}(0.57) = \begin{bmatrix} -\frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

We thereby get

$$\begin{aligned} \widehat{\mathbf{P}}(0.50, 0.75) &= \begin{bmatrix} 1 - \frac{1}{5} & \frac{1}{5} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 - \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{4}{5} & \frac{1}{5} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \underline{\underline{\begin{bmatrix} \frac{3}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}}. \end{aligned}$$