Department of Mathematical Sciences

Examination paper for TMA4275 Lifetime analysis

Examination date: June 7th 2022

Examination time (from-to): 09.00-13.00

Permitted examination support material: A / All support material is allowed

Academic contact during examination: Håkon Tjelmeland
Phone: 4822 1896

Technical support during examination: Orakel support services
Phone: 73 59 16 00

If you experience technical problems during the exam, contact Orakel support services as soon as possible before the examination time expires/the test closes. If you don't get through immediately, hold the line until your call is answered.

OTHER INFORMATION

Do not open Inspera in multiple tabs, or log in on multiple devices, simultaneously. This may lead to errors in saving/submitting your answer.

Get an overview of the question set before you start answering the questions.

Read the questions carefully, make your own assumptions and specify them in your answer. Only contact academic contact if you think there are errors or insufficiencies in the question set.

Cheating/Plagiarism: The exam is an individual, independent work. Examination aids are permitted, but make sure you follow any instructions regarding citations. During the exam it is not permitted to communicate with others about the exam questions or distribute drafts for solutions. Such communication is regarded as cheating. All submitted answers will be subject to plagiarism control. *Read more about cheating and plagiarism here.*

Citations: If using expressions from the course textbook you should refer to the equation numbers in the textbook, explain why this expression is valid in the situation considered in your problem, and start by copying the expression in your solution.

Notifications: If there is a need to send a message to the candidates during the exam (e.g. if there is an error in the question set), this will be done by sending a notification in Inspera. A dialogue box will appear. You can re-read the notification by clicking the bell icon in the top right-hand corner of the screen. All candidates will also receive an SMS to ensure that nobody misses out on important information. Please keep your phone available during the exam.

Weighting: The exam consists of several items (Problems 1, 2, 3a, 3b, 4 etc.). Each of the items is given the same weight in the evaluation of the exam.

ABOUT SUBMISSION

Answering in Inspera: In each of the problems you should upload one pdf file with your solution of that problem. If a problem consists of several items (a, b, etc.) your solutions to all these items should be included in the same pdf file. You are advised **not** to answer any problems directly in Inspera.

File upload: When working in other programs because parts of/the entire answer should be uploaded as a file attachment – make sure to save your work regularly.

All files must be uploaded before the examination time expires.

All uploaded files should be pdf files.

30 minutes are added to the examination time to manage the sketches/calculations/files. The additional time is included in the remaining examination time shown in the top left-hand corner.

NB! You are responsible to ensure that the file(s) are correct and not corrupt/damaged. Check the file(s) you have uploaded by clicking “Download” when viewing the question. All files can be removed or replaced as long as the test is open.

How to digitize your sketches/calculations
How to create PDF documents
Remove personal information from the file(s) you want to upload

Automatic submission: Your answer will be submitted automatically when the examination time expires and the test closes, as long as you have answered at least one question. This will happen even if you do not click “Submit and return to dashboard” on the last page of the question set. You can reopen and edit your answer as long as the test is open. If no questions are answered by the time the examination time expires, your answer will not be submitted. This is considered as “did not attend the exam”.

Withdrawing from the exam: If you become ill during the exam or wish to submit a blank answer/withdraw from the exam for another reason, go to the menu in the top right-hand corner and click “Submit blank”. This cannot be undone, even if the test is still open.

Accessing your answer post-submission: You will find your answer in Archive when the examination time has expired.

1 **Introduction**: The following table contains some observed life times and (right) censored times. The right censored observations are marked with an asterix ($\star$).

$$1.11, 1.13^{\star}, 1.35, 1.40^{\star}, 1.83^{\star}, 2.16, 2.22, 2.25^{\star}, 2.40, 2.65^{\star}$$

Problem: Use the above data to compute by hand the Nelson-Aalen estimate for the cumulative hazard rate. In your solution include enough numbers so that it is clear what calculations you have done. For evaluation of sums, products etc. you may use a calculator or a computer.

Based on the numbers you found for the Nelson-Aalen estimate, write your own code in R to compute the estimate for the variance of the Nelson-Aalen estimator and the associated 95% confidence interval for the cumulative hazard rate based on a log-transformation. Make also a plot showing the estimated Nelson-Aalen estimate for $t \in [0, 3]$ together with the confidence interval. In your solution include the R code you have used to produce the results.

Note: You should upload one pdf file that contains image(s) of your handwritten solution to the problem. When correcting the exam emphasis will be placed on the answers being logical and that they contain all natural intermediate calculations.

Upload your file here. Maximum one file.

All file types are allowed. Maximum file size is **50 GB**

Select file to upload

Maximum marks: 10

2 **Introduction**: Let T be a lifetime for which the hazard rate is given as

$$\alpha(t) = \lambda e^{\beta t} \text{ for } t \geq 0,$$

where $\lambda > 0$ and $\beta \neq 0$ are parameters.

Problem: Derive formulas for the survival function $S(t)$ and the density function $f(t)$ for T.

Note: You should upload one pdf file that contains image(s) of your handwritten solution to the problem. When correcting the exam emphasis will be placed on the answers being logical and that they contain all natural intermediate calculations.

Upload your file here. Maximum one file.

All file types are allowed. Maximum file size is **50 GB**

Select file to upload

Maximum marks: 10

3 **Introduction**: The R output below shows some results when estimating a Cox model to a survival data set of 241 individuals with monoclonal gammopathy of undetermined significance (MGUS). The fitted model has three covariates: age, sex and mspike. The covariate sex is binary, whereas the covariates age and mpsike are both continuous. In the data set the values of the covariate age varies from 34 to 90, and the values for mspike varies from 0.3 to 3.2.

R commands and R output:

```
> res.cox = coxph(Surv(futime,death) ~ age + sex + mspike,data=mgus)
> summary(res.cox)
Call:
coxph(formula = Surv(futime, death) ~ age + sex + mspike, data = mgus)

  n= 241, number of events= 225

              coef     exp(coef)  se(coef)      z      Pr(>|z|)
age        0.066748  1.069026  0.007051   9.466  <2e-16 ***
sexmale  0.236694  1.267053  0.136816   1.730  0.0836 .
mspike   -0.084901  0.918603  0.170947  -0.497   0.6194
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

           exp(coef)  exp(-coef)  lower .95 upper .95
age         1.0690      0.9354     1.0544     1.084
sexmale   1.2671      0.7892     0.9690     1.657
mspike     0.9186      1.0886     0.6571     1.284

Concordance= 0.689  (se = 0.019 )
Likelihood ratio test = 104  on 3 df,   p=<2e-16
Wald test            = 96.48  on 3 df,   p=<2e-16
Score (logrank) test = 100.2  on 3 df,   p=<2e-16

> res.cox$var
             [,1]          [,2]          [,3]
[1,]  4.972337e-05 -7.378061e-05 4.857322e-05
[2,] -7.378061e-05  1.871872e-02 1.957005e-04
[3,]  4.857322e-05  1.957005e-04 2.922280e-02
```

Problem: Write down the estimated relative risk function.

Find a 95% confidence interval for the ratio of the hazard rate for a female of age 50 years over the hazard rate of a female of age 51 years, when the mspike value is the same.

Find a 95% confidence interval for the relative risk function for a male of age 35 years and with mspike value equal to 1.5.

Note: You should upload one pdf file that contains image(s) of your handwritten solution to the problem. When correcting the exam emphasis will be placed on the answers being logical and that they contain all natural intermediate calculations.

Upload your file here. Maximum one file.

All file types are allowed. Maximum file size is **50 GB**

Select file to upload

Maximum marks: 10

4 **Introduction**: Consider a situation where we observe failures occurring in n units of a particular type. Assume we have no censoring and that whenever a unit is failing it is immediately repaired. We assume that a unit that has been repaired may fail again and that the intensity for failure is uninfluenced by previous failures/repairs. We assume there is no limit on how many times a unit can fail, so the number of units at risk is n at all times.

For each individual we have one (time invariant) binary covariate, $x_i \in \{0, 1\}$ for unit number i. We let the intensity process for failure of unit i be given by

$$\lambda_i(t) = \alpha_0(t)e^{\beta x_i} \text{ for } t \geq 0,$$

where $\alpha_0(t)$ is an unspecified baseline hazard rate and β is a scalar parameter.

Problem: Assuming we observe the process up to a time τ, find a formula for the partial likelihood function for β, and for the corresponding log-partial likelihood function.

Derive a simple expression for the maximum partial likelihood estimator for β.

Hint: Introduce necessary notation so that you can write the partial likelihood function and the estimator for β in a simple form.

Note: You should upload one pdf file that contains image(s) of your handwritten solution to the problem. When correcting the exam emphasis will be placed on the answers being logical and that they contain all natural intermediate calculations.

Upload your file here. Maximum one file.

All file types are allowed. Maximum file size is **50 GB**

Select file to upload

Maximum marks: 10

5 **Introduction (to part A)**: Let Z be exponentially distributed with mean $1/\lambda$, so that the density function for Z is

$$f(z) = \lambda e^{-\lambda z} \text{ for } z \geq 0.$$

Problem A: From the definition of the Laplace transform, show that the Laplace transform of Z, $\mathscr{L}(c)$, is given by

$$\mathscr{L}(c) = \frac{\lambda}{\lambda+c}.$$

Show by induction that the r'th derivative of $\mathscr{L}(c)$ is given by

$$\mathscr{L}^{(r)}(c) = (-1)^r \frac{\lambda \cdot r!}{(\lambda+c)^{r+1}}.$$

Introduction (to part B): Assume we have m clusters of individuals. Numbering the clusters from 1 to m, let n_i denote the number of individuals in cluster number i. Assume that all the n_i individuals in cluster number i share a frailty variable Z_i and that the hazard rate for each individual in this cluster is given by

$$\alpha(t|Z_i) = Z_i \cdot t^{k-1},$$

where $k > 0$ is a parameter. Finally we assume $Z_1, \ldots, Z_m$ to be independent and to be exponentially distributed with mean $1/\lambda$. Note that the values of the two parameters λ and k are the same for all individuals.

Consider a situation with right censoring, so that for individual number j in cluster number i we observe $(\widetilde{T}_{ij}, D_{ij})$, where D_{ij} is an indicator specifying whether or not the lifetime of this individual is censored and $\widetilde{T}_{ij}$ is the corresponding observed lifetime or censoring time.

Problem B: Find an expression for the log-likelihood function in terms of the quantities specified above and $D_{i\bullet} = \sum_{j=1}^{n_i} D_{ij}$.

If possible, find explicit expressions for the maximum likelihood estimators for λ and k. If this is not possible, optimise (if possible) analytically with respect one of the parameters, find the profile likelihood for the other parameter, and explain how the profile likelihood can be used to find the maximum likelihood estimates for λ and k. If it is neither possible to optimise analytically with respect to one of the parameters, discuss briefly how one can then find the maximum likelihood estimates.

Note: You should upload one pdf file that contains image(s) of your handwritten solution to both Problem A and Problem B. When correcting the exam emphasis will be placed on the answers being logical and that they contain all natural intermediate calculations.

Upload your file here. Maximum one file.

All file types are allowed. Maximum file size is **50 GB**

Select file to upload

Maximum marks: 20

6 **Introduction**: Let $\{Z_i\}_{i=0}^{\infty}$ be a discrete time Markov chain where $Z_i \in \{0,1\}$ for all $i = 1,2,\ldots$ and with $Z_0 = 0$. Let the transition probabilities be given by $P(Z_{i+1} = 1 | Z_i = 0) = \alpha$ and $P(Z_{i+1} = 0 | Z_i = 1) = \beta$ for all $i = 0,1,2,\ldots$.

Let $\{N(t); t \geq 0\}$ be a counting process which is independent of $\{Z_i\}_{i=1}^{\infty}$, and let $\lambda(t)$ denote the intensity process of $N(t)$. For $k = 1,2,\ldots$ we let T_k denote the k'th event time in $N(t)$, and set as usual $T_0 = 0$. Define a process $\{X(t), t \geq 0\}$ by

$$X(t) = \sum_{i=1}^{N(t)} Z_i \text{ for } t \geq 0.$$

Let $\{\mathscr{F}_t\}$ be the history that at time t contains all information about the counting process $N(t)$ up to and including time t and information about $Z_i, i = 1,\ldots,N(t)$.

Problem: Explain why it is obvious that $X(t)$ is a sub-martingale.

Find an expression for the incremental process $dX(t)$, and use this to derive an expression for the incremental compensator process

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

Use the expression you found for $dX^\star(t)$ to derive an expression for the compensator $X^\star(t)$.

Note: You should upload one pdf file that contains image(s) of your handwritten solution to the problem. When correcting the exam emphasis will be placed on the answers being logical and that they contain all natural intermediate calculations.

Upload your file here. Maximum one file.

All file types are allowed. Maximum file size is **50 GB**

Select file to upload

Maximum marks: 10