STOCKHOLMS UNIVERSITET
DEPT. OF MATHEMATICS
DIV. OF MATHEMATICAL STATISTICS

MT7052
EXAM
October 25 2024

Exam in Inference and prediction for life and health processes 7.5 ECTS Credits

October 25 2024

Examination by: Mathias Lindholm, `lindholm@math.su.se`.

Allowed to use: Three (3) sheets of (physical) A4 papers with notes (in any fontsize or handwritten and on both sides of each page) that are brought to the exam by the student. Pocket calculator provided by the department in connection with the exam.

Return of exam: No later than three weeks after the written exam.

Each correctly solved problem yields 10 points and the maximum total score is 50 points. All reasoning and notation must be clear. Note that vague comments tend to be interpreted as lack of understanding.

It is OK to answer in Swedish.

The grades A–E are set according to the following lower point limits:

Grade	A	B	C	D	E
Lower limit	45	40	35	30	25

Problem 1

Let $A^*(t) := \int_0^t I(Y(s) > 0)\alpha(s)\mathrm{d}s$ and introduce

$$\widehat{A}(t) := \int_0^t \frac{I(Y(s) > 0)}{Y(s)}\mathrm{d}N(s). \tag{1}$$

a) Suppose that you have observed data up until time t_0. Express the Nelson-Aalen estimator (1) for a time point $t < t_0$ in terms of data.

b) Show that $\widehat{A}(t)$ is an unbiased estimator of $A^*(t)$. Clearly state your assumptions.

c) Show that $A^*(t)$ is a biased estimator of $A(t) = \int_0^t \alpha(s)\mathrm{d}s$ and determine whether $A^*(t)$ produces an over or underestimation.

(10p)

Problem 2

Consider a population of animals that can die of two different causes, cause 1, and cause 2. We want to analyse the cause death processes, based on the following right-censored data (study time scale):

T_j	no. of animals at risk	cause of death
0.451	15	2
0.463	13	1
0.510	10	2
0.562	9	2
0.602	7	1

where T_j denotes the time of the jth observed death time. The data is subject to right-censoring and does not contain any tied events and all causes of death are assumed to occur independently.

a) Define the intensity process for the counting processes $N_1(t)$, $N_2(t)$, and $N(t)$, that counts deaths of cause 1, of cause 2, and all deaths, respectively, when assuming a multiplicative intensity model structure.

b) Estimate the cumulative hazard rate function for deaths of cause 1 and cause 2.

c) How does the estimate of cumulative hazard rates for cause specific mortality relate to the estimate of cumulative hazard rates that do not take cause into consideration?

(10p)

Problem 3

Let $\alpha(t; x) = \alpha_0(t)e^{\beta x}$, where $x = 0$ or 1. Further, let T_i denote the ith recorded time, let $\delta_i = 1$ if the ith recorded time is an observed event and 0 if the ith recorded time is right censored and finally let x_i denote the covariate coding of the ith individual.

Given the above you will record tuples (T_i, δ_i, x_i) for all individuals $i = 1, \ldots, n$. We will now consider the situation with the following data:

$$\{(T_i, \delta_i, x_i)\}_{i=1,\ldots,5} = \{(1, 1, 1), (3, 1, 0), (4, 0, 1), (7, 0, 0), (10, 1, 1)\}.$$

a) Write out the risk set for each time point where an event is observed.

b) Set up Cox's partial likelihood for the data above and find the ML-estimate for β given data.

c) Carry out a likelihood ratio test for the hypothesis that $\beta = 0$ at the 95%-level. What is the alternative hypothesis?

Hint: recall that $\chi^2_{LR} = 2(\log L(\widehat{\beta}) - \log L(\beta_0)) \sim \chi^2(1)$, where $\chi^2(1)_{0.025} = 5.02$ and $\chi^2(1)_{0.05} = 3.84$.

(10p)

Problem 4

Let $\alpha(t \mid Z) = \alpha(t)Z$, and assume that Z follows a Gamma-distribution with

$$\mathcal{L}_Z(s) = \mathrm{E}[\exp\{-sZ\}] = (1 + \delta s)^{-1/\delta}.$$

a) Show that the population hazard rate $\mu(t) = -S'(t)/S(t)$ and calculate $\mu(t)$ in the above setting.

b) Let $\alpha(t) = \exp\{\beta t\}$ and determine the limit of $\mu(t)$ as $t \to +\infty$, and interpret the result.

(10p)

Problem 5

Consider a continuous time Markov chain $Z(t)$ that can only take on the values 0, 1, or 2, where 0 corresponds to alive and healthy, 1 corresponds to sick, and 2 that corresponds to dead. The only transition rates $\alpha_{ij}(t)$ that are non-zero, are $\alpha_{01}(t), \alpha_{02}(t), \alpha_{10}(t)$, and $\alpha_{12}(t)$.

a) Draw a graph that describes the possible states with allowed transitions indicated by arrows.

b) Let $P_{ij}(s,t) := \mathrm{P}(Z(t) = j \mid Z(s) = i)$. The Kolmogorov forward equations (KFE) are given by

$$\frac{\partial}{\partial t} P_{ij}(s,t) = \sum_{k:k\neq j} P_{ik}(s,t)\alpha_{kj}(t) - P_{ij}(s,t)\alpha_j(t),$$

where $P_{ij}(s,s) = \delta_{ij}$, $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, and where

$$\alpha_j(t) := \sum_{k:k\neq j} \alpha_{jk}(t).$$

For the described process $Z(t)$, express the KFE for the relevant transition probabilities $P_{ij}(s,t)$ and simplify.

c) Draw an example trajectory of $Z(t)$ with two jumps over a time interval $[0, t_0]$, when $Z(0) = 1$.

(10p)

Good luck!