Imperial College London

MATH70048

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May 2023

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science

Survival Models

Date: 15 May 2023

Time: 14:00 – 16:30 (BST)

Time Allowed: 2.5hrs

This paper has 5 Questions.

Please Answer Questions 1-2, and Questions 3-5 in Separate Answer Booklets

Candidates should start their answers to each question on a new sheet of paper.

Supplementary books may only be used after the relevant main book(s) are full.

Any required additional material(s) will be provided.

Allow margins for marking.

Credit will be given for all questions attempted.

Each question carries equal weight.

DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO

© 2023 Imperial College London

1. (a) State the definition of the hazard rate.

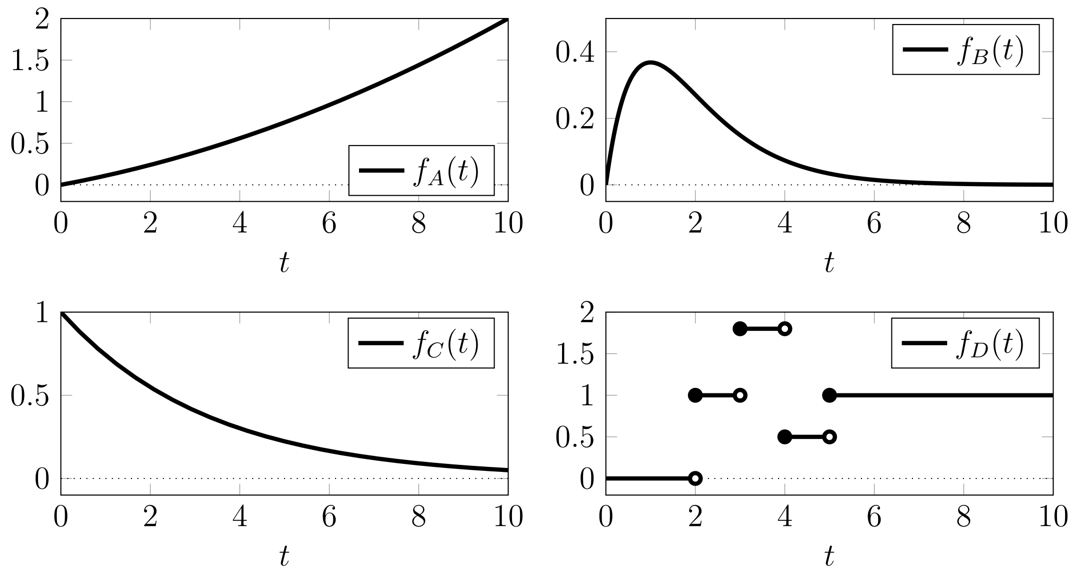
Use this definition to derive the hazard function for a random variable $T \sim$ Exponential(λ) for some $\lambda > 0$.

You may use $P(T > x) = \exp(-\lambda x)$ for $x > 0$ and the memoryless property of the exponential distribution, but no other properties. (6 marks)

(b) Let $T^{(1)}, T^{(2)}, \dots, T^{(n)}$ be independent nonnegative continuous random variables with hazard functions $h^{(1)}(t), \dots, h^{(n)}(t)$.

Derive the hazard rate of $T = \min(T^{(1)}, \dots, T^{(n)})$. (4 marks)

(c) Below you find graphs of four functions $f_A(t)$, $f_B(t)$, $f_C(t)$ and $f_D(t)$.



State which of these functions could be

(i) a survivor function,

(ii) a hazard rate,

(iii) an integrated hazard rate,

(iv) a cumulative distribution function,

(v) a probability density function?

Justify your answers, including whether you need to check additional properties that you cannot immediately read off from this graph.

(10 marks)

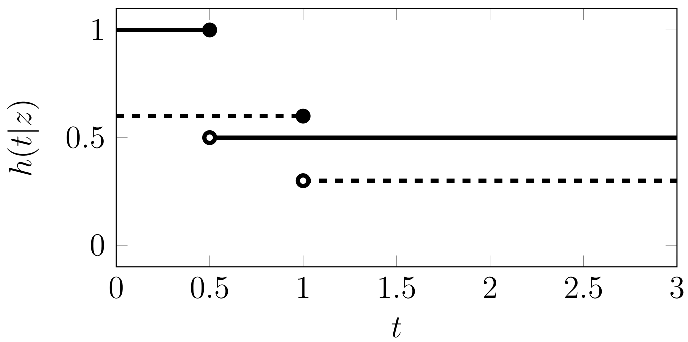
(Total: 20 marks)

2. (a) Define Cox's proportional hazard model via the individual hazard rates, describing all variables. (3 marks)

(b) Describe why it is called a proportional hazards model. (2 marks)

(c) Give the general form of the partial likelihood for such a model, describing all variables. (3 marks)

(d) The following plot gives the hazard rates $h(t|z)$ for two different individuals (dotted line and dashed line). Could these arise from a proportional hazards model? Explain your answer.



(2 marks)

(e) Suppose we observe 6 individuals that are at risk to some event, each with one covariate x, with possible right censoring as described in the following table.

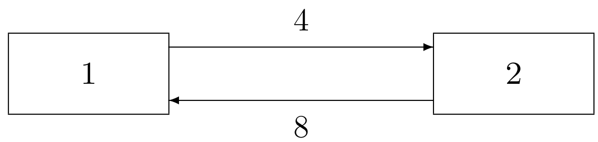
at risk time	1	2	3	5	6	11
event observed	no	no	yes	no	yes	yes
covariate x	2	2	4	1	5	1

(i) Give the partial likelihood function in this example (without simplifying it).
Based on this expression, describe briefly how you would compute an estimator of the unknown parameter. (4 marks)

(ii) We would like to investigate whether x has a statistically significant effect on the lifetime. We decide to use a likelihood ratio test.
Perform this test and comment on the appropriateness of this approach in this situation.
The following information might be useful. The maximised log-likelihood under the model is -1.234.
$\log(6) = 1.792$, $\log(7) = 1.946$, $\log(8) = 2.079$, $\log(9) = 2.197$, $\log(10) = 2.303$, $\log(11) = 2.398$, $\log(12) = 2.485$
If $X \sim \chi^2_1$ *then* $P(X > 2.706) \approx 0.9$ *and* $P(X > 3.841) \approx 0.95$.
If $X \sim \chi^2_2$ *then* $P(X > 4.605) \approx 0.9$ *and* $P(X > 5.991) \approx 0.95$.
If $X \sim \chi^2_3$ *then* $P(X > 6.251) \approx 0.9$ *and* $P(X > 7.815) \approx 0.95$.
(6 marks)

(Total: 20 marks)

3. Consider a 2-state Markov model with transition intensities as per the following diagram. We assume that we move through it via a continuous Markov jump process $X(t)$.



(a) Which states are absorbing? (2 marks)

(b) Write down the generator matrix G.

(3 marks)

(c) Derive the transition probabilities $p^{21}(t)$, where the transition probabiltiy $p^{21}(t)$ is as defined in the lectures.

You may use

$$\frac{dy}{dt} = ay + b \quad \Longrightarrow \quad y = ke^{at} - \frac{b}{a},$$

for $a \neq 0, b, k \in \mathbb{R}$. (6 marks)

(d) Suppose we start in state 2 at time $t = 0$.
What is the long-run probability of being in state 2? (2 marks)

(e) State and prove the Chapman-Kolmogorov equations for a general time-homogeneous n-state Markov jump process $X(t)$. Introduce any notation that you are using. (7 marks)

(Total: 20 marks)

4. (a) Define a martingale with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$. (3 marks)

(b) Consider a nonnegative random variable T with hazard rate

$$h(t) = \begin{cases} 2 & \text{for } t \leq 1 \\ 1 & \text{for } t > 1. \end{cases}$$

Define the associated counting process $N(t)$ and the at-risk indicator based on observing one individual with a potential right-censoring time C in terms of T and C.

State the intensity of $N(t)$ for the above $h(t)$.

Suppose we observe one realisation of this setup, an uncensored observation at time $t = 2$. Sketch the counting process, its compensator with respect to the filtration generated by itself and the resulting martingale process on the interval $[0, 3]$.

(7 marks)

(c) Give the definition of a Poisson process. (3 marks)

(d) Discuss briefly whether Poisson processes are suitable models for the classical survival analysis setting, where lifetimes of individuals are being observed that are subject to right-censoring. (2 marks)

(e) Consider a Poisson process N defined on $[0, 2]$ with intensity

$$\lambda(t) = \begin{cases} \theta & \text{for } t \leq 1 \\ 2\theta & \text{for } t > 1 \end{cases}, \quad t \in [0, 2].$$

Suppose we want to estimate the parameter θ using maximum likelihood estimation.

Write down a suitable likelihood and derive the maximum likelihood estimator.

Suppose we observe a realisation of the Poisson process with 4 events which occur at the time points 0.1, 1.2, 1.5, 1.9. Compute the maximum likelihood estimator for this realisation.

(5 marks)

(Total: 20 marks)

5. This question is based on the introduction to Section 4.2, and on Sections 4.2.1, 4.2.2, 4.2.4, 8.1 of

Aalen, Borgan, and Gjessing. 2008. Survival and Event History Analysis: A Process Point of View. Springer.

(a) Define the additive model proposed by Aalen in terms of the hazard rate, stating what the parameters of this model are.

Is this model parametric, semi-parametric or nonparametric? (3 marks)

(b) Give one disadvantage and one advantage of this additive hazard model compared to the Cox model. (4 marks)

(c) Write the additive Aalen model using counting processes, integrated regression functions and martingales in matrix form, defining all notation.

State how the elements of this form would be called in a linear model.

(6 marks)

(d) State two differences compared to a typical linear regression model.

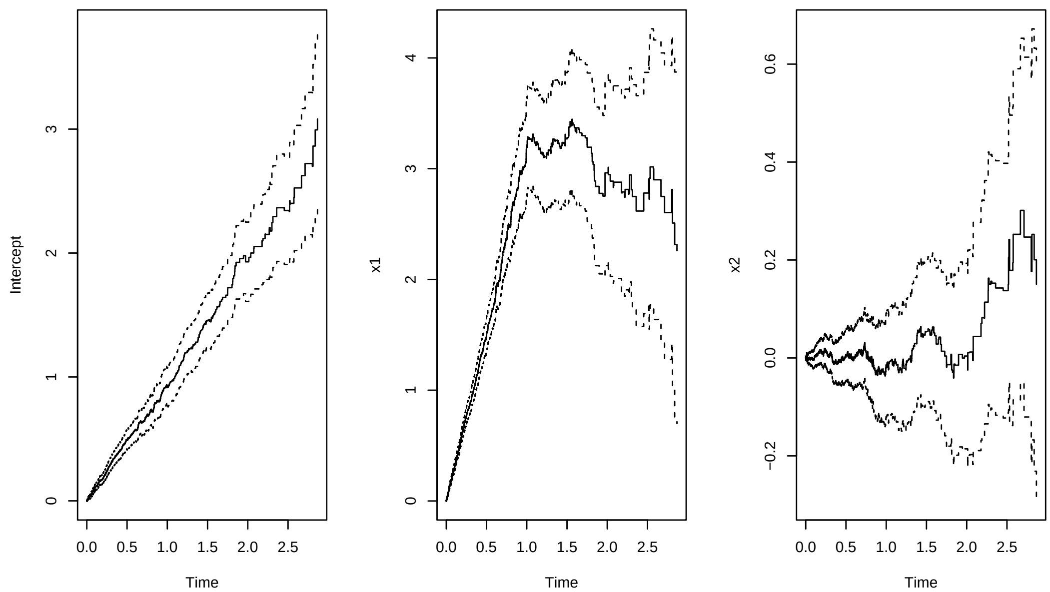
(2 marks)

(The question continues on the next page)

(e) Consider an additive Aalen model with 2 covariates x_1 and x_2. Below is a plot of the integrated effect estimates of the intercept and these two covariates, together with 95% confidence intervals.

Which variables seem to have a statistically significant effect?

Describe the effect on the hazard rate of the intercept and of any significant covariates, stating roughly what the corresponding coefficient could be.



(3 marks)

(f) Suppose we observe n individuals and model when they eat during a day. Suppose we decide to model for each individual a counting process throughout the day with Aalen's additive regression model. As covariates we use an intercept, the age and the weight of the individuals. Can martingale theory be used to analyse results in this study? (2 marks)

(Total: 20 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2023

This paper is also taken for the relevant examination for the Associateship.

MATH60048/70048

Survival Models (Solutions)

Setter's signature	Checker's signature	Editor's signature
....	....	....

1. (a) *(the case $\lambda = 1$ was computed in the lectures)* The hazard function of T is defined as a function $h : [0, \infty) \mapsto [0, \infty)$, with $\boxed{\text{sim. seen} \Downarrow}$

$$h(x) = \lim_{\Delta \downarrow 0} \frac{P(T \le x + \Delta | T > x)}{\Delta}.$$

Suppose now that $T \sim$ Exponential(λ) then

$$h(x) = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \frac{1 - \exp(-\lambda(x + \Delta)) - (1 - \exp(-\lambda x))}{\exp(-\lambda x)} = \lim_{\Delta \downarrow 0} \frac{1 - \exp(-\lambda \Delta)}{\Delta}$$

(candidates are allowed to use the memoryless property of the exponential distribution to skip the intermediate step) Using l'Hôpital's rule, we get

$$h(x) = \lim_{\Delta \downarrow 0} \lambda \exp(-\lambda \Delta)/1 = \lambda.$$

$\boxed{6, \text{B}}$

(b) Let $S^{(j)}(t)$ be the survivor function of $T^{(j)}$, so $S^{(j)}(t) = \exp(-\int_0^t h^{(j)}(s) ds)$. $\boxed{\text{seen} \Downarrow}$
If $T = \min(T^{(1)}, T^{(2)}, \dots, T^{(n)})$ then $\forall t \ge 0$, $\{T \ge t\} = \bigcap_{j=1}^n \{T^{(j)} \ge t\}$. Hence, by independence of the $\{T_j\}$,

$$S(t) = P(T \ge t) = P\left(\bigcap_{j=1}^n \{T^{(j)} \ge t\}\right) = \prod_{j=1}^n P(T^{(j)} \ge t) = \prod_{j=1}^n S^{(j)}(t)$$
$$= \prod_{j=1}^n \exp\left\{-\int_0^t h^{(j)}(s) ds\right\} = \exp\left\{-\int_0^t \sum_{j=1}^n h^{(j)}(s) ds\right\}.$$

By the identity $S(t) = \exp\{-\int_0^t h(s) ds\}$ and the Fundamental Theorem of Calculus this implies

$$h(t) = \sum_{j=1}^n h^{(j)}(t).$$

$\boxed{4, \text{B}}$

(c) (i) Only function f_C can be a survivor function. It is the only one that is non-increasing (candidates can also say decreasing), and it has $0 \le f_C(t) \le 1$. $\boxed{\text{unseen} \Downarrow}$ $\boxed{2, \text{B}}$

(ii) All four functions could be hazard rates. $\boxed{2, \text{A}}$

(iii) Only function $f_A(t)$ can be an integrated hazard rate, as all others are not non-decreasing (*increasing is also acceptable*). $\boxed{2, \text{B}}$

(iv) None of the functions can be a cumulative distribution function.
Only f_A is non-decreasing, but $f_A(10) > 1$, which eliminates the possibility of it being a cumulative distribution function. $\boxed{2, \text{B}}$

(v) Only $f_B(t)$ could be a probability density function.
The other functions integrate to more than 1. For $f_B(t)$ the integral could be 1, but this would have to be verified as it cannot be directly read from the plot. $\boxed{2, \text{C}}$

2. (a) The Cox model proposes that the hazard rates of individuals are related via the relationship

$$h(t; z) = h_0(t) \exp(\beta \cdot z),$$

where $\beta \in \mathbb{R}^p$ is a vector of regression parameters, $z \in \mathbb{R}^p$ is a p-dimensional covariate, $\beta \cdot z = \sum_{i=1}^{p} \beta_i z_i$ is the inner product and $h_0(t)$ is the baseline hazard, the hazard of a (possibly hypothetical) individual with $z = 0$.

seen ⇓

3, A

(b) Under the Cox model, the hazard functions of two individuals with covariates z_1, z_2 are in constant proportion at all times; that is,

seen ⇓

$$\frac{h(t; z_1)}{h(t; z_2)} = \frac{\exp(\beta \cdot z_1)}{\exp(\beta \cdot z_2)} = \exp\{\beta \cdot (z_1 - z_2)\},$$

giving rise to the name *proportional hazards*.

2, A

(c)

$$L(\beta) = \prod_{i \in U} \frac{\exp(\beta \cdot z_i)}{\sum_{j \in R_{t_i}} \exp(\beta \cdot z_j)},$$

where the product is over the set U of *uncensored* observations (death times) and t_i are the corresponding event times.

R_{t_i} is the *risk set*; that is, the set of individuals "at risk" (or "in view", so not dead or censored) at time t_i.

3, A

(d) No. Proportionality of the hazards is not given. Consider, for example, the periods $[0, 0.5]$ and $(0.5, 1]$ where the ratio between the hazards of these individuals is clearly different.

unseen ⇓

An alternative answer would state that the hazard rates "jump" at different times, which is not possible in a proportional hazards model.

2, D

(e) (i)

meth seen ⇓

$$L(\beta) = \frac{\exp(4\beta)}{\exp(4\beta) + \exp(\beta) + \exp(5\beta) + \exp(\beta)} \cdot \frac{\exp(5\beta)}{\exp(5\beta) + \exp(\beta)}$$

Numerical optimisation can be used to derive the maximum likelihood estimator.

4, A

(ii) The maximised partial likelihood under the null (with $\beta = 0$) is $L(0) = 1/4 \cdot 1/2 = 1/8$.

The likelihood ratio statistic is

$$W = 2(\log(L(\hat{\beta})) - \log(L(0))) = 2(-1.234 + 2.079) = 2 * 0.845 = 1.690.$$

As there is one degree of freedom, this needs to be compared to the quantiles of the χ_1^2 distribution. The test statistic is less than both the 90% and the 95% quantile, implying that the null hypothesis of no effect would not be rejected at the 10% and 5% level.

This approach is not appropriate, as the asymptotic theory cannot be relied upon in a data set this small. [*only 3/6 marks can be reached in this question if this is not clearly stated*].

6, D

3. (a) No state is absorbing. sim. seen ⇓

2, A

(b)

$$G = \begin{pmatrix} -4 & 4 \\ 8 & -8 \end{pmatrix}$$

3, A

(c) Using the Kolmogorov forward equations,

$$\begin{aligned}\frac{d}{dt}p^{21}(t) &= -p^{21}(t)\cdot 4 + p^{22}(t)\cdot 8 = -p^{21}(t)\cdot 4 + (1-p^{21}(t))\cdot 8 \\ &= 8 - 12p^{21}(t).\end{aligned}$$

From the hint, this gives

$$p^{21}(t) = k\exp\{-12t\} + \frac{8}{12} = k\exp\{-12t\} + \frac{2}{3}$$

for some $k \in \mathbb{R}$. Using the boundary condition $p^{21}(0) = 0$ yields $k = -2/3$ and hence

$$p^{21}(t) = \frac{2}{3}(1 - \exp(-12t)).$$

6, C

(d) As we are starting in state 2, we are looking for the limit of $p^{22}(t)$ as $t \to \infty$:

$$\lim_{t\to\infty} p^{22}(t) = 1 - \lim_{t\to\infty} p^{21}(t) = 1 - \frac{2}{3} = \frac{1}{3}.$$

2, B

(e) The Chapman-Kolmogorov equations are seen ⇓

$$\mathbf{P}_{s+t} = \mathbf{P}_s\mathbf{P}_t, \quad \text{if } s,t \geq 0,$$

where $\mathbf{P}_t = (p^{ij}(t))_{i,j=1,\dots,n}$ and $p^{ij}(t) = P(X(t) = j|X(0) = i)$. 3, A

Thus,

$$\begin{aligned}(\mathbf{P}_{s+t})_{ij} = p^{ij}(s+t) &= P(X(s+t) = j|X(0) = i) \\ &= \sum_{k=1}^{n} P(X(s) = k|X(0) = i)P(X(s+t) = j|X(s) = k) \\ &= \sum_{k=1}^{n} p^{ik}(s)p^{kj}(t) = (\mathbf{P}_s\mathbf{P}_t)_{ij}.\end{aligned}$$

4, C

4. (a) A stochastic process $\{X(t)\}$ is a *martingale* with respect to the filtration $\mathcal{F}_t$ if, $\forall t$, $\boxed{\text{seen} \Downarrow}$

1. $\mathsf{E}[|X(t)|] < \infty$;
2. $\mathsf{E}[X(t)|\mathcal{F}_s] = X_s \ \forall t > s$.

$\boxed{3, \text{A}}$

(b) The counting process is

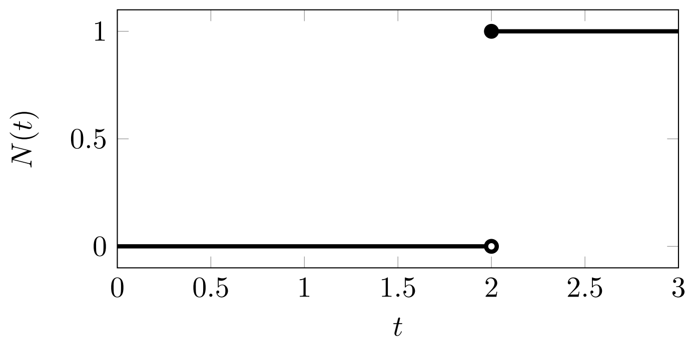
$$N(t) = I(T \leq t, T \leq C), \quad t > 0,$$

where $I(A)$ is the indicator function of the event A.
The at-risk indicator is

$$Y(t) = I(T \geq t, C \geq t), \quad t > 0.$$

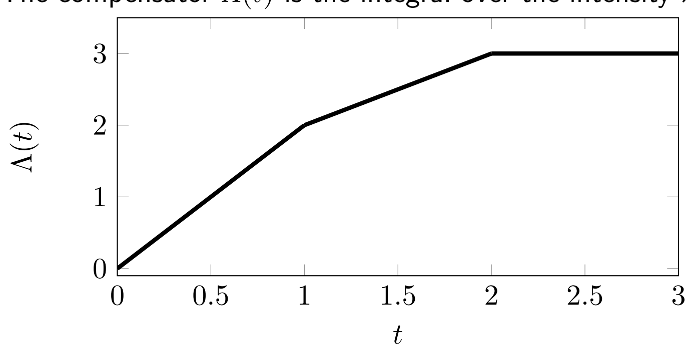
As derived in the lectures, the intensity process is $\lambda(t) = Y(t)h(t)$, which translates to $\lambda(t) = 2Y(t)$ for $t \leq 1$ and $\lambda(t) = Y(t)$ for $t > 1$. $\boxed{3, \text{A}}$

The counting process for this realisation is as follows: $\boxed{\text{unseen} \Downarrow}$



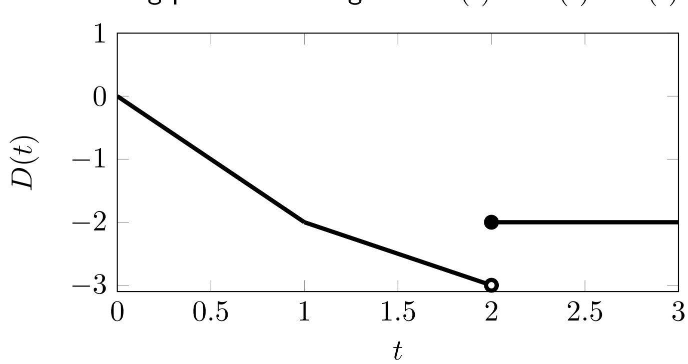
$\boxed{1, \text{A}}$

The compensator $\Lambda(t)$ is the integral over the intensity $\lambda(t)$.



$\boxed{2, \text{B}}$

The counting process martingale is $D(t) = N(t) - \Lambda(t)$.



$\boxed{1, \text{D}}$

(c) A Poisson process is a counting process $\{N(t) : t \geq 0\}$ with independent increments. seen $\Downarrow$

3, A

(d) Poisson processes are not suitable for modeling right-censored observations. This is because they have a deterministic intensity. This does not allow adjusting the intensity according to the at-risk status (i.e. the intensity of counting processes of events should be able to change once individuals get right-censored or have an event). unseen $\Downarrow$

2, D

(e) In the lecture we have seen that the number of events observed in a Poisson process over a fixed time interval follows a Poisson distribution with expectation being equal to the cumulative intensity over the time period. unseen $\Downarrow$

Here, the integrated intensity is $\Lambda(2) = \int_0^2 \lambda(s)ds = \theta + 2\theta = 3\theta$.

Thus, letting j denote the number of observed events, a likelihood is

$$L(\theta) = \exp(-3\theta)\frac{(3\theta)^j}{j!}.$$

The corresponding log-likelihood is

$$\log L(\theta) = -3\theta + j\log(3\theta) - \log(j!)$$

with derivative

$$\frac{\partial}{\partial\theta}\log L(\theta) = -3 + j/\theta.$$

Equating this to 0 and checking the second derivative, which is negative, shows that the MLE is indeed $\hat{\theta} = j/3$.

For the particular realisation, this equates to $\hat{\theta} = 4/3$.

5, D

5. (a) The hazard rate at time t for an individual i with vector of covariates $x_i(t) = (x_{i1}(t), \ldots, x_{ip}(t))^T$ takes the form

seen ⇓

$$\alpha(t|x_i) = \beta_0(t) + \beta_1(t)x_{i1}(t) + \ldots + \beta_p(t)x_{ip}(t)$$

The parameters are the functions $\beta_0(t), \ldots, \beta_p(t)$. The model is nonparametric. 3, M

(b) One disadvantage of the additive hazard model is that it does not automatically force hazards to be non-negative.

The material lists several possible advantages (on p.155 and 156 of the material). Students can name any one of them, e.g.:

* An additive model might be a more correct representation of the relation between the covariates and the hazard rates in a situation. Particularly the linear dependence of the hazard rate on parameter functions $x_{ij}(t)$ might be more appropriate.
* Additivity (or linearity) fits nicely with the underlying martingale theory. Estimates and residuals give exact martingales, which is not the case for the Cox model. 4, M

(c) In counting process notation, we work with counting processes $N_i(t)$ that has an intensity

$$\lambda_i(t) = Y_i(t)(\beta_0(t) + \beta_1(t)x_{i1}(t) + \ldots + \beta_p x_{ip}(t)),$$

where $Y_i(t)$ is the at risk indicator.

Let $N(t) = (N_1(t), \ldots, N_n(t))^T$ be the vector of observed counting processes.

Let $M(t) = (M_1(t), \ldots, M_n(t))^T$ be the corresponding vector of counting process martingales.

The vector of cumulative regression functions is $B(t) = (B_0(t), \ldots, B_p(t))^T$, where $B_j(t) = \int_0^t \beta_j(s)ds$.

Let $X(t)$ be the $n \times (p+1)$ matrix with i-th row $(Y_i(t), Y_i(t)x_{i1}(t), \ldots, Y_i(t)x_{ip}(t))$.

Then the model can be written in differential form as

$$dN(t) = X(t)dB(t) + dM(t).$$

This has the same form as a linear model, with $dN(t)$ taking the role of the observation vector, $X(t)$ taking the role of the design matrix, $dB(t)$ taking the role of the unknown parameter vector and $dM(t)$ taking the role of the error term. 4, M

(d) (*this has not been discussed explicitly in the underlying material, so candidates can give different answers)*

unseen ⇓

The error term in a linear regression model is often assumed to follow a normal distribution. In the Aalen model, the error distribution being is a counting process martingale, and would not be expected to be resembling a normal distribution - it could be skewed or bimodal. (*on p. 169 of the material a similar thought is mentioned in relation to residuals).*

For each individual we have one regression model for each time point -in typical linear regression models there is only one observation per individual. 4, M

(e) The intercept and the covariate x_1 seem to have a significant effect, as 0 is not contained in the confidence bands. unseen $\Downarrow$

x_2 does not seem to have an effect as 0 is contained in the confidence interval throughout.

The estimate of $B_0(t)$ for the intercept seems to have a relatively constant slope of close to 1, thus the corresponding coefficient could be in the region of 1 throughout.

The estimate of $B_1(t)$ seems to be linearly increasing to about 3 by time 1. Afterwards there seems to be no clear significant change. Thus the corresponding coefficient would be in the region of 3 until time 1, but then changes to something not significant (e.g. 0). 3, M

(f) No. This model does not take information about previous events of the individual into account. Thus it is only a rate model, and not an intensity model. unseen $\Downarrow$

2, M

Review of mark distribution:

Total A marks: 32 of 32 marks
Total B marks: 20 of 20 marks
Total C marks: 12 of 12 marks
Total D marks: 16 of 16 marks
Total marks: 100 of 100 marks (including Mastery)
Total Mastery marks: 20 of 20 marks

If your module is taught across multiple year levels, you might have received this form for each level of the module. You are only required to fill this out once for each question.

ExamModuleCode	QuestionNumber	Comments for Students
MATH60048/70048	1	This question was generally well answered. In part (c), the markers put emphasis on justifications, not only for those functions that could be functions of the desired type, but also for those functions that could not. Several candidates incorrectly stated that probability density functions (pdfs) must be less than or equal to 1.
MATH60048/70048	2	The question was generally answered well. In Part (a) and (c) candidates sometimes did provide answers that did not clearly define all variables that are involved. In part e (i), some candidates did not discuss the need for numerical methods. In part e (ii) many candidates forget to discuss the lack of appropriateness of the method in this setting due to small sample sizes. There was a typo in e(ii) in the exam: in the last three lines, where quantiles of the chi-squared distribution are given, all inequality signs should be < instead of >. This was announced during the exam; judging from the answers given by students, it did not lead to confusion or increase the difficulty of the exam.
MATH60048/70048	3	Most students performed well on this question. The majority got full marks in parts (a) and (b). Many students also got full marks for part (c), though a common error here was to solve the ODE for p21(t) by treating p22(t) as a constant, instead of first substituting p22(t)=1-p21(t). Part (d) was achieved by most, with many using the straightforward approach in the solutions. Several students gave an argument that showed that p22=1/3 was the corresponding component of the chain's stationary distribution, without definitively linking this to the limiting distribution; this was awarded half-marks. Part (e) was also reasonably well done by most, though many students lost one or two marks due to a lack of detail in their presented proofs, or for attributing the wrong steps of their proof to the Markov property / time homogeneity.
MATH60048/70048	4	Q4 proved challenging for most members of the cohort. In part (a), most students lost at least one mark for errors in their mathematical notation; for example, in giving the finite expectation property as a conditional expectation, given the filtration. For part (b), several students failed to give the counting process, at-risk indicator and intensity process in terms of T & C, as requested, though the majority of students managed to get most marks available for the sketches. Several students lost marks in part (c), mostly for a lack of sufficient clarity in the definition. Most students got at least half marks for part (d); full marks here required demonstration of the understanding that in the classical survival analysis setting, the current intensity depends on the history of the process (past events / past censoring). Part (e) was generally poorly done, with most students using an incorrect approach to constructing the likelihood.

MATH70048	5	Q5 was reasonably well done. Most students obtained the correct form of the hazard rate model in part (a), though a significant number of students failed to adequately indicate the parameters of the model. Part (b) was generally successfully completed. For part (c), most students managed to get the correct matrix form of the additive Aalen model, but a large proportion lost at least one mark for a lack of clarity/detail in defining their notation. Part (d) was often well done, with many students picking appropriate differences between the additive Aalen model and a generic linear regression model. In general, part (e) proved a little more challenging - most students correctly identified the intercept and the first covariate as having statistically significant effect, but not the second covariate; unfortunately, a large number of students failed to give any statement of (roughly) what the corresponding coefficients would be, losing marks in the process. Part (f) was correctly answered by about half of the cohort.