IMPERIAL

MATH70048

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May 2024

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

Survival Models

Date: Monday, May 13, 2024

Time: 10:00 – 12:30 (BST)

Time Allowed: 2.5 hours

This paper has 5 Questions.

Please Answer All Questions in 1 Answer Booklet

Candidates should start their solutions 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

1. Let N_t denote the number of failures/renewals in the interval $(0, t)$, where immediately after each failure, the process renews with an independent renewal increment T_i specifying the time between the $(i-1)$th and ith failure. The distributions of $T_1, T_2, \ldots$ are identical and specified by density functions $f(t)$ at $t \in \mathbb{R}^+$ with Laplace transform $f^*(s)$ at $s \in \mathbb{C}$.

(a) With $Z_r = T_1 + T_2 + \cdots + T_r$ the time to the rth renewal, write the probability mass function of N_t in terms of the distribution function of Z_r. (3 marks)

(b) The probability generating function for N_t is defined in the usual way as

$$G(t, \zeta) = \sum_{r=0}^{\infty} \zeta^r \text{pr}(N_t = r)$$

and it can be show that the Laplace transform of this probability generating function is of the form

$$G_o^*(s, \zeta) = \frac{1 - f^*(s)}{s\{1 - \zeta f^*(s)\}}.$$

Suppose that the distribution of each renewal increment T_i is exponential of rate ρ. Show that $f^*(s) = \rho/(\rho + s)$ and thereby obtain a more explicit expression for $G_o^*(s, \zeta)$.

(2 marks)

(c) Suppose now that the interval over which the event process is observed is $(0, T)$, where T is a random variable having an exponential distribution of rate λ. Obtain an expression for the unconditional probability generating function for the number of renewals in $(0, T)$ in the form

$$G(\zeta) = \frac{\theta}{1 - (1 - \theta)\zeta} \qquad (1)$$

for suitably defined θ. (6 marks)

(d) It can be shown that random variables with distribution characterised by $G(\zeta)$ in equation (1) have mean $1/\theta$. Either from the physical description of the problem, or based on the answer to part (c), specify the role of λ relative to ρ in determining the expected number of renewals in the random interval $(0, T)$, where T is as in part (c). (3 marks)

(e) Use the physical interpretation of the special Erlangian distribution of rate λ and index β and the information provided in part (d) to deduce the expected number of renewals in $(0, T)$ when the renewal increments are exponentially distributed of rate ρ as above, and T is special Erlangian of rate λ and index β.

(6 marks)

(Total: 20 marks)

2. (a) Let $T_1, \ldots, T_m$ be independent and identically distributed random variables from a Weibull distribution with rate parameter ρ and shape parameter α. The associated survivor function is

$$S(t) = \exp(-\rho^\alpha t^\alpha).$$

(i) Prove that $Y = \min\{T_1, \ldots, T_m\}$ has a Weibull distribution with rate parameter $\theta = m^{1/\alpha}\rho$ and shape parameter α. (2 marks)

(ii) Specify the distributions of $T_1^\alpha, \ldots, T_m^\alpha$ and compare the distribution of $\min\{T_1^\alpha, \ldots, T_m^\alpha\}$ to that of Y^α, where Y is defined above. (3 marks)

(b) Suppose that for each of n individuals, a survival time Y_i is observed, where $i = 1, \ldots, n$. One generating process for each Y_i is as the minimum survival time from m competing sources of failure. With the survival time for each source of failure specified as a Weibull distribution of rate ρ_i and shape α, Y_i has a Weibull distribution of rate $\theta_i = m^{1/\alpha}\rho_i$ and shape α, as derived above.

(i) By introducing a vector of covariates $x \in \mathbb{R}^p$, whose values for individual i are denoted by x_i, one possible model expresses θ_i directly in terms of x_i as $\theta_i = \exp(x_i^{\mathrm{T}}\beta_\theta)$, while a second model instead models $\rho_i = \exp(x_i^{\mathrm{T}}\beta_\rho)$. Are the two models compatible? (3 marks)

(ii) Discuss the interpretation of the parameters β_θ and β_ρ of the models constructed in part (b)(i). (4 marks)

(iii) Show that the Weibull model constructed above with θ_i modelled as $\theta_i = \exp(x_i^{\mathrm{T}}\beta_\theta)$ belongs to the accelerated life family. (3 marks)

(iv) Show that the same Weibull model belongs to the proportional hazards family. (5 marks)

(Total: 20 marks)

3. (a) Define the terms *nuisance parameter* and *interest parameter*. (2 marks)

(b) Let (Y_{i0}, Y_{i1}) be exponentially distributed random variables of rates ρ_i/ψ and $\rho_i\psi$ respectively, where both $\psi > 0$ and $\rho_i > 0$ are fixed parameters (i.e. not treated as random variables), and Y_{i0}, Y_{i1} are independent.

(i) Derive the density function of $Z_i = Y_{i1}/Y_{i0}$ and confirm that this does not depend on ρ_i. (4 marks)

(ii) Explain how the observation in part (b)(i) can be used as a basis for inference from realisations of n independent pairs of random variables (Y_{i0}, Y_{i1}), $i = 1, \ldots, n$. (2 marks)

(iii) If the observation from part (b)(i) was not used, what complications would arise concerning inference on ψ? (4 marks)

(iv) Suppose now that the above exponential model is modified such that (Y_{i0}, Y_{i1}) are Weibull distributed random variables with common shape parameter α, assumed known, and rate parameters ρ_i/ψ and $\rho_i\psi$ respectively. Identify the transformation of (Y_{i0}, Y_{i1}) that eliminates ρ_i. (6 marks)

(v) Explain, in not more than 6 sentences, the construction of a matched-pair design for assessing a single treatment. With reference to an example of your choice, explain what is encapsulated within the parameters ρ_i and ψ in the context of the exponential model of part (b)(i). (2 marks)

(Total: 20 marks)

4. (a) (i) Without equations, explain the meaning of the terms *sufficient statistic* and *minimal sufficient statistic* in the context of a given parametric model parametrised by θ. (2 marks)

(ii) Show, using the example of an exponential family, how a sufficient statistic of lower dimension than n can sometimes be deduced from the log-likelihood function corresponding to a sample of size n. Note that a density function is said to be of exponential family form, parametrised in terms of its canonical statistic θ if it can be written as

$$f(y;\theta) = \exp\{o(y)^{\mathrm{T}}\theta - \kappa(\theta)\} f_0(y)$$

for known functions $o(y)$ and $\kappa(\theta)$. (4 marks)

(b) (i) With θ partitioned as (ψ, λ), where ψ is the interest parameter, indicate the general form of a likelihood function that allows λ to be eliminated from the inference problem for ψ via a partial likelihood. (2 marks)

(ii) Specialise the partial likelihood construction above to one based on a conditional density function, and one based on a marginal density function, starting from the density function $f_{(S,R)}(s, r; \psi, \lambda)$ of a jointly sufficient statistic (S, R). (3 marks)

(c) (i) Consider data consisting of realisations of $Y_1, \ldots, Y_n$ and $D_1, \ldots, D_n$, where $Y_i = \min\{T_i, c_i\}$, c_i is a non-random censoring time, T_i a survival time, and D_i indicates whether or not individual i is censored. Specifically, $D_i = 1$ if $Y_i = T_i$, and otherwise $D_i = 0$. Let $f_T(t;\psi)$ and $S_T(t;\psi)$ be the density and survivor functions for each T_i. Construct the likelihood function for ψ in terms of f_T and S_T and express it only in terms of the hazard function h_T. You do not need to make any modelling assumptions at this point. (4 marks)

(ii) Suppose that the effects of covariates are modelled in the form of a proportional hazards model $h(t;x) = h_0(t) g(x^{\mathrm{T}}\beta)$ where $g(x^{\mathrm{T}}\beta)$ is known but the baseline hazard function $h_0(t)$ is unspecified. Survival times $T_1, \ldots, T_n$ can be treated as conditionally independent given the covariates. Indicate the difficulties with the formulation in part (c)(i). (2 marks)

(iii) It can be shown that the likelihood function for the proportional hazards model with censored observations factorises as

$$L(\beta, h_0; \mathscr{D}) = \prod_{j\in\mathcal{U}} \left(\frac{g(x_{i_j}^{\mathrm{T}}\beta)}{\sum_{k\in\mathscr{R}_j} g(x_k^{\mathrm{T}}\beta)} \right) \prod_{j\in\mathcal{U}} w(y_{(1)}, \ldots, y_{(j)}; h_0, \beta) \prod_{j\in\mathcal{C}} S_T(y_{(j)}; \beta, h_0),$$

for some function $w(\cdot)$, where $\mathscr{D}$ represents all the available data, $\mathcal{U}$ and $\mathcal{C}$ are the index sets for the uncensored and censored observations respectively, $\mathscr{R}_j$ is the index set for the individuals still at risk at the time of the jth ordered survival time $y_{(j)}$, and i_j is the index of the individual who fails at time $y_{(j)}$. Briefly discuss how β can be estimated. (3 marks)

(Total: 20 marks)

5. This question is based on De Stavola, B. and Cox, D. R. (2017). "Detecting bias arising from delayed recording of time", *J. R. Statist. Soc. Ser. C*, 66, 1065–1073.

Three random variables play a key role: V (unobserved) is the time between the origin and the event; Z (unobserved) is the time between the origin and entry into the study; T (directly observed) is the time between entry into the study and the event.

(a) (i) Explain with a suitable diagram the identity $T = (V - Z)^+$. What practical situation might lead $V - Z$ to be negative? (2 marks)

(ii) Explain in not more than 4 sentences why the objective of the study is the dependence of V on covariates rather than the dependence of T on covariates. (1 mark)

(b) The measured dependence of T on covariates x is partitioned into two parts: the dependence of pr$(T = 0)$ on x and the dependence of the distribution of T on x conditional on $T > 0$. From this conditioning operation, there results a strictly positive random variable T^+. Suppose that V and Z are independent and each has an exponential distribution with rate parameters ρ_V and ρ_Z respectively.

(i) Show that the hazard functions associated with V and Z are ρ_V and ρ_Z respectively at all evaluation points. (2 marks)

(ii) Show that

$$\text{pr}(T = 0) = \frac{\rho_V}{\rho_V + \rho_Z}.$$

(5 marks)

(iii) Show algebraically that the distribution of T^+ is the same as that of V. (6 marks)

(iv) What is the practical implication of part (b)(iii)? (2 marks)

(c) Suppose now that dependence of the distributions of V and Z is via a proportional hazards representation with

$$\begin{aligned} \rho_V &= \rho_{V_0} \exp\{\beta_V^{\mathrm{T}} x\} \\ \rho_Z &= \rho_{Z_0} \exp\{\beta_Z^{\mathrm{T}} x\}, \end{aligned}$$

where ρ_{V_0} and ρ_{Z_0} are the rate parameters determining the probabilistic behaviour of V and Z when the covariates are at baseline $x = 0$. Show that

$$\log\left\{\frac{\text{pr}(T = 0)}{\text{pr}(T > 0)}\right\} = \log(\rho_{V_0}) - \log(\rho_{Z_0}) + (\beta_V - \beta_Z)^{\mathrm{T}} x.$$

(2 marks)

(Total: 20 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2024

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) Let N_t be the number of renewals in the interval $(0, t)$ and let $Z_r = T_1 + \cdots + T_r$ be the time to the rth renewal. Then $N_t < r$ if and only if $Z_r > t$. Thus, with $F_{Z_r}(t)$ the distribution function of Z_r, seen ⇓ 3, A

$$\text{pr}(N_t < r) = \text{pr}(Z_r > t) = 1 - F_{Z_r}(t),$$

and with $F_{Z_0}(t) = 1$,

$$\text{pr}(N_t = r) = \text{pr}(N_t < r + 1) - \text{pr}(N_t < r) = F_{Z_r}(t) - F_{Z_{r+1}}(t).$$

seen/sim.seen ⇓

(b) By the definition of the Laplace transform 2, A

$$f^*(s) = \rho \int_0^\infty e^{-st} e^{-\rho t} dt = \frac{\rho}{\rho + s}.$$

On substituting this into the expression for $G_o^*(s, \zeta)$ we obtain

$$G_o^*(s, \zeta) = \frac{1}{\rho + s - \zeta\rho}.$$

meth seen ⇓

(c) From the lectures we know that if the distribution of T is special Erlangian of rate λ in $\beta \in \mathbb{Z}^+$ stages, then the unconditional probability generating function can be obtained from the Laplace transform of the conditional probability generating function through differentiation with respect to s as 6, C

$$G(\zeta) = \frac{\lambda^\beta}{\Gamma(\beta)} \left(-\frac{\partial}{\partial s}\right)^{\beta - 1} G^*(s, \zeta)\bigg|_{s=\lambda},$$

which provides a way to calculate $G(\zeta)$ without inversion of the Laplace transform. With $\beta = 1$ this becomes

$$G_o(\zeta) = \frac{\lambda}{(\rho + s - \zeta\rho)}\bigg|_{s=\lambda} = \frac{\lambda}{\rho + \lambda - \zeta\rho},$$

which is of the form

$$G(\zeta) = \frac{\theta}{1 - (1 - \theta)\zeta}$$

with $\theta = \lambda/(\rho + \lambda)$.

unseen ⇓

(d) From the expression in part (iii), $1/\theta = 1 + \rho/\lambda$, tending to 1 and ∞ as $\rho/\lambda \to 0$ and $\rho/\lambda \to \infty$ respectively. This accords with intuition as $1/\lambda$ is the expected length of the interval and ρ is the rate at which events occur along the interval. An alternate version of this answer could refer to the scaling trade off between ρ and λ, e.g. doubling both leaves the expected value unchanged. 3, B

unseen ⇓

(e) Since a special Erlangian distribution can be viewed as the sum of the times spent in β stages of life, where the times in each stage are exponential of rate λ, T can be written as $T = Y_1 + \cdots + Y_\beta$, where each Y_j is exponentially distributed of rate λ. Within any single stage, the expected number of failures is $1/\theta$ as indicated in the the statement of part (iii), and by the properties of expectations, the expected number of failures over the β stages is therefore β/θ.

An alternative solution modifies the calculation of part (iii) above to show that the distribution of the number of renewals is negative binomial of index β and parameter θ. From this, the form of the negative binomial probability mass function can be used to deduce the expected number of renewals. The relevant calculation for the expectation must be provided to obtain full marks. 6, D

seen/sim.seen ⇓

2. (a) (i) If the minimum exceeds t, then all m exceed t, thus 2, A

$$\text{pr}(Y > t) = \exp\{-m\rho^\alpha t^\alpha\} = \exp\{-\theta^\alpha t^\alpha\},$$

the survivor function of a Weibull random variable of rate θ and shape α.

seen/sim.seen ⇓

(ii) For each i, the survivor function of T_i is 3, B

$$\text{pr}(T_i > t) = \exp\{-m\rho^\alpha t^\alpha\}.$$

Thus

$$\text{pr}(T_i^\alpha > t) = \text{pr}(T_i > t^{1/\alpha}) = \exp(-\rho^\alpha t), \qquad (1)$$

which is the survivor function of an exponential random variable of rate ρ^α. By the same argument, Y^α has an exponential distribution of rate θ^α. Consider on the other hand, the survivor function of the minimum of m random variables with survivor function (1). This is $\exp(-m\rho^\alpha t)$. It follows that $\min\{T_1^\alpha, \ldots, T_m^\alpha\}$ has the same distribution as $(\min\{T_1, \ldots, T_m\})^\alpha$, namely, exponential of rate θ^α.

unseen ⇓

(b) (i) By the definition of θ_i, compatibility of the two models entails 3, B

$$\exp(x_i^{\mathrm{T}}\beta_\theta) = m^{1/\alpha}\exp(x_i^{\mathrm{T}}\beta_\rho) = \exp\{x_i^{\mathrm{T}}\beta_\rho + \log(m^{1/\alpha})\}$$

and since there is no x_i-free value of β_θ that equalises the left and right-hand sides, the two models are in contradiction unless x_i includes 1 as its first entry. In this case, any discrepancy between the two models can be absorbed in the intercept term, the effect of covariates on the distribution of survival times being stable.

seen/sim.seen ⇓

(ii) Consider two covariate vectors x and $\tilde{x}$, say, which coincide exactly except for in the last entries x_p and $\tilde{x}_p$ which differ by one unit. Then 4, A

$$\frac{\exp(x^{\mathrm{T}}\beta_\theta)}{\exp(\tilde{x}^{\mathrm{T}}\beta_\theta)} = \exp\{(x - \tilde{x})^{\mathrm{T}}\beta_\theta\} = \exp(\pm\beta_{\theta,p}).$$

Thus, two hypothetical individuals with covariates x and $\tilde{x}$ have survival times with proportional rates $\exp(x^{\mathrm{T}}\beta_\theta) = \exp(\pm\beta_{\theta,p})\exp(\tilde{x}^{\mathrm{T}}\beta_\theta)$, where $\beta_{\theta,p}$ is the pth entry of β_θ.

This proportionality of rates also holds in the second of the two models due to cancellation of the $m^{1/\alpha}$ term. In particular, $\exp(x^{\mathrm{T}}\beta_\rho) = \exp(\pm\beta_{\rho,p})\exp(\tilde{x}^{\mathrm{T}}\beta_\rho)$.

meth seen ⇓

3, A

(iii) The accelerated life model is defined by the condition

$$S(y;x) = S_0\{a(x)y\}, \tag{2}$$

for some function $a : \mathbb{R}^p \to \mathbb{R}^+$ such that $a(0) = 1$. The survivor functions of a Weibull random variable of rate θ and shape α is

$$S(t) = \exp(-\theta^\alpha t^\alpha),$$

and if $\theta = \theta(x) = \exp(x^{\mathrm{T}}\beta_\theta)$ then $\theta(0) = 1$, verifying equation (2).

meth seen ⇓

5, B

(iv) The proportional hazards model is defined by the condition

$$h(y;x) = h_0(y)g(x), \tag{3}$$

for some function $g : \mathbb{R}^p \to \mathbb{R}^+$ such that $g(0) = 1$.
The hazard function in terms of the survivor function is

$$h(t) = -\frac{d}{dt}\log S(t) = \alpha\theta^\alpha t^{\alpha-1}.$$

thus if $\theta = \theta(x) = \exp(x^{\mathrm{T}}\beta_\theta)$, $\theta(0) = 1$ and, on taking $g(x) = \theta(x)^\alpha$, the proportional hazards assumption (3) is satisfied.

3. (a) An interest parameter is one of core scientific or societal interest, typically with a subject-matter interpretation. A nuisance parameter is only needed to complete the specification of the probabilistic model, enabling inference on the interest parameter. seen ⇓ 2, A

(b) (i) In the following calculation we drop the subscript i. Since the exponential survivor function is slightly simpler in form than the distribution function, consider seen ⇓ 4, B

$$\begin{aligned}\text{pr}(Z > z) &= \text{pr}(Y_1 > zY_0) = (\rho/\psi)\int_0^\infty \text{pr}(Y_1 > zy_0)\exp\{-\rho y_0/\psi\}dy_0\\ &= (\rho/\psi)\int_0^\infty \exp\{-\rho y_0(\psi z + 1/\psi)\}dy_0 = \frac{1}{1+z\psi^2}.\end{aligned}$$

Differentiation of the resulting distribution function shows that the density function of Z at z is

$$f_Z(z) = \frac{\psi^2}{(1+\psi^2 z)^2}, \tag{4}$$

which is free of the nuisance parameter ρ.

(ii) From realisations of $Z_1, \ldots, Z_n$ each having density function (4), a log-likelihood can be constructed as seen ⇓ 2, A

$$\ell(\psi; z_1, \ldots, z_n) = \sum_{i=1}^n \log\Big\{\frac{\psi^2}{(1+\psi^2 z_i)^2}\Big\}$$

and maximised with respect to ψ. Inference for ψ follows by standard likelihood theory.

(iii) Without the preliminary manoeuvres described above, a standard likelihood analysis would involve a very large parameter space, growing proportionally to the number of observations in the sample. Standard likelihood theory entails large-sample approximations, but these apply under a notional asymptotic regime in the sample size when the dimension of the parameter space is fixed. If each pair contributes a parameter, standard asymptotic approximations to the distributions of Wald, Rao score and likelihood ratio statistics do not straightforwardly apply and there is no immediate way to calibrate the inference. unseen ⇓ 4, D

(iv) As in the solution to part (b)(i), the i subscript is suppressed in the following calculations. unseen ⇓ 6, D

The Weibull survivor function is for Y_1 is

$$\text{pr}(Y_1 > y_1) = \exp\{-(\rho\psi y_1)^\alpha\}$$

It follows that

$$\text{pr}(Y_1^\alpha > q_1) = \text{pr}(Y_1 > q_1^{1/\alpha}) = \exp\{-(\rho\psi)^\alpha q_1\}$$

which is the survivor function of an exponential random variable of rate $(\rho\psi)^\alpha$. An analogous calculation shows that Y_0^α has the distribution of an exponential random variable of rate $(\rho/\psi)^\alpha$. By the appropriate modification of the calculation in part (b)(i), we obtain

$$\begin{aligned}\text{pr}(Z^\alpha > z) &= (\rho/\psi)^\alpha \int_0^\infty \exp\{-\rho^\alpha q_0(z\psi^\alpha + 1/\psi^\alpha)\} dq_0 \\ &= \frac{\rho^\alpha}{\psi^\alpha \rho^\alpha (z\psi^\alpha + 1/\psi^\alpha)}\end{aligned}$$

The density function of Z^α is therefore

$$f_{Z^\alpha}(z) = \frac{\psi^{2\alpha}}{(1 + z\psi^{2\alpha})^2},$$

which is free of ρ.

There is a second possible solution that does not require α to be known but is harder to see. The Weibull distribution is in the scale family so the ratio $Z_i = Y_{i1}/Y_{i0}$ eliminates ρ_i here too, and the resulting density function at z is

$$\frac{\alpha\psi^2 z^{\alpha-1}}{(1 + \psi^2 z^\alpha)^2}.$$

A suitable solution would need to derive the previous displayed equation.

$\boxed{\text{seen/sim.seen} \Downarrow}$

(v) A paired-comparison study entails randomising a treatment to one unit in each pair and leaving the other as the untreated control. The parameter ρ_i is a pair-specific nuisance parameter representing, for instance, genetic differences between pairs of twins, likely to effect the outcome, but not of direct concern for understanding the effect of the treatment, which is encapsulated in ψ. Both parameters enter the model by modifying the rate parameter of the exponential random variable.

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

4. (a) (i) A sufficient statistic in the context of a given parametric model parametrised by θ is one that contains all the information in the sample relevant for inference on θ. Since the data themselves are sufficient, as is the likelihood function, interest lies in the statistic that achieves the most effective compression, i.e. is of lowest dimension. Such a statistic is called a minimal sufficient statistic.

seen ⇓

2, A

(ii) The likelihood function corresponding to n independent observations from an exponential family takes the form

seen ⇓

4, A

$$\prod_{i=1}^{n} f(y_i;\theta) = \exp\Big\{\sum_{i=1}^{n} o(y_i)^{\mathrm{T}}\theta - n\kappa(\theta)\Big\}\prod_{i=1}^{n} f_0(y_i),$$

viewed as a function of θ, with $y_1,\ldots,y_n$ fixed at their observed values. Since constants of proportionality are irrelevant for inference on θ, the term $\prod f_0(y_i)$ can be ignored, showing that $\sum_{i=1}^{n} o(y_i)$ is sufficient for θ. Provided that the dimension of θ is substantially lower than n, an appreciable reduction in dimension can be achieved by retaining only $\sum_{i=1}^{n} o(y_i)$ rather than the original data $y_1,\ldots,y_n$.

(b) (i) The general form is

seen ⇓

2, A

$$L(\psi,\lambda;y) = L_{\mathsf{pa}}(\psi;y)L_{\mathsf{r}}(\psi,\lambda;y).$$

where the partial likelihood $L_{\mathsf{pa}}(\psi;y)$ does not depend on λ and is used for inference on ψ, and the remainder likelihood $L_{\mathsf{r}}(\psi,\lambda;y)$ is discarded.

(ii) Since (S,R) is jointly sufficient for (ψ,λ) inference on these parameters can be based on the joint density function of (S,R) which decomposes as

seen ⇓

3, A

$$f_{S,R}(s,r;\psi,\lambda) = f_S(s;\psi,\lambda)f_{R|S}(r\,|\,s;\psi,\lambda),$$

A useful conditional likelihood can be constructed in the special case that $f_{R|S}(r\,|\,s;\psi,\lambda) = f_{R|S}(r\,|\,s;\psi)$, i.e., the conditional density function is free of λ. In that case the partial likelihood is taken as

$$L_{\mathsf{pa}}(\psi;y) = f_{R|S}(r\,|\,s;\psi)$$

Two types of useful marginal likelihood are available either when $f_S(s;\psi,\lambda) = f_S(s;\psi)$ is free of λ or when this marginal density function factorises as $f_S(s;\psi,\lambda) = g(s;\psi)h(s;\psi,\lambda)$. In these cases the partial likelihood is either $L_{\mathsf{pa}}(\psi;y) = f_S(s;\psi)$ or $L_{\mathsf{pa}}(\psi;y) = g(s;\psi)$.

(c) (i) The general form of the likelihood from density and survivor functions $f_T(t;\psi)$ and $S_T(t;\psi)$ of T at t is

seen ⇓

4, C

$$L(\psi;\mathscr{D}) = \prod_{i=1}^{n} f_T(y_i;\psi)^{d_i} S_T(y_i;\psi)^{1-d_i}, \tag{5}$$

where $\mathscr{D}$ is the data $\{y_1,\ldots,y_n,d_1,\ldots,d_n\}$. Let $\mathcal{U} \triangleq \{i : d_i = 1\}$ denote the uncensored observations, let $\mathcal{C} \triangleq \{i : d_i = 0\}$ denote the censored observations, and set $r(u) = |\{i : y_i \geq u\}|$. After taking logarithms in (5)

$$
\begin{aligned}
\ell(\psi;\mathscr{D}) &= \sum_{i\in\mathcal{U}} \log f(y_i;\psi) + \sum_{i\in\mathcal{C}} \log S(y_i;\psi) \\
&= \sum_{i\in\mathcal{U}} \log h(y_i;\psi) + \sum_{i=1}^{n} \log S(y_i;\psi) \\
&= \sum_{i\in\mathcal{U}} \log h(y_i;\psi) - \sum_{i=1}^{n} \int_0^{y_i} h(u;\psi)du \\
&= \sum_{i\in\mathcal{U}} \log h(y_i;\psi) - \int_0^{\infty} r(u)h(u;\psi)du, \qquad (6)
\end{aligned}
$$

where the integrand in the last expression is zero beyond the last observed survival or censoring time.

unseen $\Downarrow$

2, B

(ii) If the dependence of $h(y;\psi)$ on y and ψ is left unspecified, standard likelihood theory does not apply. In effect, h is an infinite-dimensional nuisance parameter. Moreover, without simplifying assumptions for the form of the dependence of h on ψ, it is impossible to separate interest and nuisance parameters in a partial likelihood construction as given in part (c)(i).

seen $\Downarrow$

3, A

(iii) The partial likelihood from part (b)(i) is specified as

$$
L_{pa}(\beta;\mathscr{D}) = \prod_{j\in\mathcal{U}} \left(\frac{g(x_{i_j}^{\mathrm{T}}\beta)}{\sum_{k\in\mathscr{R}_j} g(x_k^{\mathrm{T}}\beta)} \right)
$$

and maximised with respect to β.

5. (a) (i) The first diagram (or first and second diagrams) from De Stavola and Cox (2017) can be reproduced. The random variable $V - Z$ is negative when the event is only recorded after it has taken place. $\boxed{2, \mathrm{M}}$

$\boxed{2, \mathrm{M}}$

(ii) The random variable Z describes the properties of the measurement process, while the physical or biological phenomenon of core scientific interest is encapsulated in V. It is therefore of primary interest from a scientific point of view to understand how the distribution of V depends on the covariates.

$\boxed{2, \mathrm{M}}$

(b) (i) The survivor function for V at v is $S_V(v) = e^{-\rho_V v}$ and the hazard function in terms of the survivor function is

$$h_V(v) = -\frac{d}{dv}\log S_V(v) = \rho_V.$$

Similarly for Z.

$\boxed{5, \mathrm{M}}$

(ii) The event $\{T = 0\}$ is equal to the event $\{V \leq Z\}$. Thus

$$\text{pr}(V \leq Z) = 1 - \int_0^\infty e^{-\rho_V z}\rho_Z e^{-\rho_Z z}dz = 1 - \rho_Z \int_0^\infty e^{-(\rho_V+\rho_Z)z} = \frac{\rho_V}{\rho_V + \rho_Z}.$$

$\boxed{6, \mathrm{M}}$

(iii) From Bayes's formula, and since $\{T > t, T > 0\} = \{T > t\}$,

$$\text{pr}(T^+ > t) = \text{pr}(T > t \mid T > 0) = \frac{\text{pr}(T > t)}{\text{pr}(T \neq 0)}$$

From part (b)(ii) we know that

$$\text{pr}(T \neq 0) = \frac{\rho_Z}{\rho_V + \rho_Z}.$$

Consider

$$\text{pr}(T > t) \;=\; \text{pr}(V - Z > t) = \rho_Z \int_0^\infty e^{-\rho_V(t+z)}e^{-\rho_Z z}dz = e^{-\rho_V t}\frac{\rho_Z}{\rho_V + \rho_Z}$$

It follows that

$$\frac{\text{pr}(T > t)}{\text{pr}(T \neq 0)} = e^{-\rho_V t} = \text{pr}(V > t).$$

$\boxed{1, \mathrm{M}}$

(v) The random variable V of core scientific interest is not directly observable. However since its distribution coincides with that of T, the latter can be used for inference on ρ_V.

$\boxed{2, \mathrm{M}}$

(c)

$$\frac{\text{pr}(T = 0)}{\text{pr}(T > 0)} = \frac{\text{pr}(T = 0)}{\text{pr}(T \neq 0)} = \frac{\rho_V}{\rho_Z}$$

and on substituting in the proportional hazards representation this becomes

$$\frac{\text{pr}(T = 0)}{\text{pr}(T > 0)} = \frac{\text{pr}(T = 0)}{\text{pr}(T \neq 0)} = \frac{\rho_{V_0}}{\rho_{Z_0}}\exp\{(\beta_V - \beta_Z)^{\mathrm{T}}x\}.$$

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 80 marks
Total Mastery marks: 20 of 20 marks