IMPERIAL

MATH70048

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
Summer 2025

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

Survival Models

Date: Wednesday, May 28, 2025

Time: Start time 14:00 – End time 16:30 (BST)

Time Allowed: 2.5 hours

This paper has 5 Questions.

Please Answer All Questions in 1 Answer Booklet

This is a closed book examination.

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.

Credit will be given for all questions attempted.

Allow margins for marking.

DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO DO SO

1. (a) What is a survival-time outcome? (1 mark)

(b) Give an example of a survival-time outcome that does not entail death or failure. (1 mark)

(c) Four equivalent ways of characterising a distribution of survival times are via the density function $f(t)$, the distribution function $F(t)$, the survivor function $S(t)$ and the hazard function $h(t)$.

(i) Using the representations

$$\begin{aligned} f(t)\Delta t &= \text{pr}(t < T \leq t + \Delta t) + o(\Delta t), & (\Delta t \to 0^+), \\ h(t)\Delta t &= \text{pr}(t < T \leq t + \Delta t \mid T > t) + o(\Delta t), & (\Delta t \to 0^+), \end{aligned}$$

explain the identity $h(t) = f(t)/S(t)$. (3 marks)

(ii) Show that $f(t)$ may be written in the form

$$f(t) = h(t) \exp\Big\{-\int_0^t h(u)du\Big\}.$$

(3 marks)

(d) How would negative ageing be reflected in the shape of the hazard function? (1 mark)

(e) Give an example where negative ageing may be expected. (2 marks)

(f) Suppose that $T_1 > 0$ and $T_2 > 0$ have exponential distributions with rates $\lambda_1 > 0$ and $\lambda_2 > 0$ respectively where $\lambda_1 \neq \lambda_2$. Find the probability density function of $T = T_1 + T_2$, where T_1 and T_2 are independent random variables.

Hints:

- The probability density function of a gamma random variable V with shape parameter $\alpha > 0$ and rate parameter $\rho > 0$ is

$$f_V(v)dv = \frac{\rho^\alpha}{\Gamma(\alpha)} v^{\alpha-1} e^{-\rho v} dv.$$

- The gamma integral is

$$\Gamma(a) := \int_0^\infty e^{-s} s^{a-1} ds.$$

- You may use the standard inverse Laplace transform

$$\mathscr{L}^{-1}\{(q-a)^{-(m+1)}; t\} = \frac{t^m e^{at}}{\Gamma(m+1)},$$

and the linearity property of inverse Laplace transforms

$$\mathscr{L}^{-1}\{ak_1^*(q) + bk_2^*(q); t\} = a\mathscr{L}^{-1}\{k_1^*(q); t\} + b\mathscr{L}^{-1}\{k_2^*(q); t\}.$$

(9 marks)

(Total: 20 marks)

2. (a) Suppose that data $y_1, \dots, y_n$ are observed. How does a statistical model relate to these data? Incorporate in your answer the words 'random variables', 'probability distribution', and 'parameter'. (6 marks)

(b) Suppose that, for a model parametrised by $\gamma \in \Gamma = \mathbb{R}^+$, the effect of a treatment is a relative increase or decrease from γ to $\gamma\psi$ where $\psi \in \mathbb{R}^+$. By verifying the group properties, show that the treatment effect is a group element $g \in \mathcal{G}$ acting on the parameter space as $\mathcal{G} \times \Gamma \to \Gamma$, $(g, \gamma) \to g\gamma$. What is the null treatment effect? (5 marks)

(c) For matched pairs of experimental units (e.g. identical twins), one receiving treatment and the other not, it is natural to treat paired outcomes $(Y_{i0}, Y_{i1})_{i=1}^n$ as coming from the same distribution with parameters γ_i and $g\gamma_i$, for $g \in \mathcal{G}$ a group element operating on the parameter space as $\mathcal{G} \times \Gamma \to \Gamma$. How could dependence on intrinsic variables such as biological sex or underlying health conditions be incorporated into such a model? (4 marks)

(d) Explain how the inclusion of intrinsic variables removes the need for pairing identical individuals. (5 marks)

(Total: 20 marks)

3. (a) A proportional hazards model postulates that an individual with covariate vector x has a hazard function $h(y;x)$ at time point y that relates to the baseline hazard function $h_0(y)$ as $h(y;x) = g(x^{\mathrm{T}}\beta)h_0(y)$. In this model, $g : \mathbb{R} \to \mathbb{R}^+$ and $h_0(y)$ is the hazard function of a notional individual with covariates at baseline, i.e. $x = 0$.

   (i) Show that the survivor and density functions associated with the proportional hazards model are of the form

$$S(y;x) = \{S_0(y)\}^{g(x^{\mathrm{T}}\beta)}, \qquad f(y;x) = g(x^{\mathrm{T}}\beta)\{S_0(y)\}^{g(x^{\mathrm{T}}\beta)-1}f_0(y),$$

   where $S_0(y)$ and $f_0(y)$ are the survivor and density functions associated with the process at baseline $x = 0$. (4 marks)

   (ii) Suppose that two individuals with covariates x_i and x_j produce outcomes obeying the proportional hazards assumption with $g(x^{\mathrm{T}}\beta) = \exp(x^{\mathrm{T}}\beta)$. How do their two hazard functions relate to one another? (2 marks)

   (iii) Assuming the last variable in the p-dimensional vector x is a binary treatment indicator, encoded as 0 and 1, what is the interpretation of β_p, the last entry of β, when $g(x^{\mathrm{T}}\beta) = \exp(x^{\mathrm{T}}\beta)$ in the proportional hazards model? (2 marks)

   (iv) The Weibull distribution with shape parameter α and rate parameter ρ has density and survivor functions

$$f(y) = \alpha\rho^{\alpha}y^{\alpha-1}\exp\{-(\rho y)^{\alpha}\}, \quad S(y) = \exp\{-(\rho y)^{\alpha}\}.$$

   Suppose that the effect of covariates x is modelled through the rate parameter as $\rho = \rho(x) = \exp(x^{\mathrm{T}}\beta)$, so that the baseline rate is $\rho_0 := \rho(0) = 1$. Show that this model belongs to the proportional hazards family. (4 marks)

(b) An accelerated life model postulates that an individual with covariate vector x has a survivor function $S(y;x)$ at time point y that relates to the baseline survivor function $S_0(y)$ as $S(y;x) = S_0(a(x^{\mathrm{T}}\beta)y)$, where $a : \mathbb{R} \to \mathbb{R}^+$ and $S_0(y)$ is the survivor function of a fictitious individual with covariates at baseline, i.e. $x = 0$.

   (i) Give a diagrammatic depiction of the accelerated life assumption, showing how an accelerator function $a(x^{\mathrm{T}}\beta)$ taking a value of 1/2 alters the baseline survivor function. (4 marks)

   (ii) Show that the model from part (a)(iv) belongs to the accelerated life family. (4 marks)

(Total: 20 marks)

4. (a) 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)

(b) Let $Y_1, \ldots, Y_n$ be independent random variables from a gamma distribution with shape parameter $\alpha > 0$, rate parameter $\rho > 0$, both unknown. The corresponding density function is provided in Question 1 (f).

(i) Show that the minimal sufficient statistic is $T = (S, C) = (\sum_{i=1}^{n} \log Y_i, \sum_{i=1}^{n} Y_i)$. (4 marks)

(ii) Show that for fixed α, C is sufficient for β. (4 marks)

(iii) Find a conditional likelihood function of α that does not depend on ρ. Hint: you may use that the sum of n independent gamma (α, ρ) random variables is itself gamma distributed with shape and rate parameters $(n\alpha, \rho)$. (4 marks)

(c) (i) Define partial likelihood, both in general, and with reference to part (b)(iii). (4 marks)

(ii) Briefly explain, without any equations, how partial likelihood is used in the analysis of the proportional hazards model. (2 marks)

(Total: 20 marks)

5. This question is based on Hougaard, P. (1984). "Life table methods for heterogeneous populations", *Biometrika*, 71, 75–83. I am using the notation Y for the survival time outcome, for compatibility with the lecture notes in the regression setting; Hougaard used T.

(a) Give the general form of Hougaard's frailty model in the absence of covariates. (2 marks)

(b) Suppose, as Hougaard does, that the conditional survivor function at y conditional on frailty $Z = z$ is given by $S_{Y|Z}(y|z) = \exp\{-zM(y)\}$. What is the relationship between the unconditional survivor function and the Laplace transform $f_Z^*(s)$ of the density function $f_Z(z)$ of the frailty variable? (2 marks)

(c) Give an expression, in terms of $f_Z^*(s)$, for the probability that Z falls in the infinitesimal region $[z, z + dz)$ conditional on survival beyond time y. (4 marks)

(d) How would you expect the distribution of frailty conditional on survival to compare to the distribution of frailty overall? (2 marks)

(e) Introduce covariates into the model in a way that recovers the proportional hazards model in the special case that the distribution of the frailty variable is degenerate, putting all its mass on a single point. (2 marks)

(f) In contrast to the proportional hazards model, which models the effects of covariates relative to baseline multiplicatively on the hazard scale, an additive hazards model models the effects of covariates additively on the hazards scale as

$$h(y; x) = h_0(y) + g(x^{\mathrm{T}}\beta), \quad g : \mathbb{R} \to \mathbb{R}^+,$$

where $h_0(y)$ is the baseline hazard function at y. A version with a frailty variable, analogous to Hougaard's model is

$$h_{Y|Z}(y \mid z; x) = h_0(y) + g(x^{\mathrm{T}}\beta) + z$$

where Z is a positive random variable.

(i) Show that the survivor function conditional on $Z = z$ is

$$S_{Y|Z}(y|z; x) = \exp\Big\{-zy - g(x^{\mathrm{T}}\beta)y - \int_0^y h_0(u)du\Big\}.$$

(2 marks)

(ii) Express the unconditional survivor function in terms of the Laplace transform of the density function $f_Z(z)$ of the frailty variable. (2 marks)

(iii) Deduce the form of the unconditional hazard function when the frailty variable has a gamma distribution with shape parameter α and rate parameter $\rho = 1/\alpha$. Hint: see question 1 (f) for the form of the gamma density function. (4 marks).

(Total: 20 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2025

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) A survival-time outcome is the time from some natural origin of time until an event of interest. seen $\Downarrow$ 1, A

(b) An example in which the event of interest is a success rather than a failure might be time from the onset of symptoms of a disease until recovery from the disease. seen/sim.seen $\Downarrow$ 1, A

(c) (i) By the definition of conditional probability seen $\Downarrow$

$$\text{pr}(t < T \leq t + \Delta t \mid T > t) = \frac{\text{pr}(t < T \leq t + \Delta t,\ T > t)}{\text{pr}(T > t)},$$

The denominator is the survivor function and since $\{t < T \leq t + \Delta t\} \subset \{T > t\}$ the numerator reduces to $\text{pr}(t < T \leq t + \Delta t)$. The conclusion follows on passing to the limit in $\Delta t \to 0$. 3, B

(ii) By the identity derived in the previous question seen $\Downarrow$

$$h(t) = -\frac{d}{dt} \log S(t),$$

so that 3, A

$$\int_0^t h(u)du = -(\log S(t) - \log S(0))$$

Since $S(0) = 1 - F(0) = 1$,

$$S(t) = \exp\left\{-\int_0^t h(u)du\right\} \tag{1}$$

and we obtain

$$f(t) = h(t)S(t) = h(t)\exp\left\{-\int_0^t h(u)du\right\}.$$

(d) Negative ageing is reflected in a decreasing hazard function. seen $\Downarrow$ 1, A

(e) Any example in which failures or successes occur relatively more frequently early on in the process. For example, the event of interest might be insurance claims measured from the time at which an individual passes their driving test. seen $\Downarrow$ 2, A

(f) By independence unseen $\Downarrow$

$$\mathbb{E}(e^{-qT}) = \mathbb{E}(e^{-qT_1})\mathbb{E}(e^{-qT_2})$$

where

$$\mathbb{E}(e^{-qT_j}) = \lambda_j \int_0^\infty e^{-t(q+\lambda_j)}dt = \frac{\lambda_j}{q + \lambda_j}$$

Thus

$$\mathbb{E}(e^{-qT_1})\mathbb{E}(e^{-qT_2}) = \frac{\lambda_1}{q+\lambda_1}\frac{\lambda_2}{q+\lambda_2}.$$

The partial-fraction expansion of the denominator is 2, C

$$\frac{1}{(q+\lambda_1)(q+\lambda_2)} = \frac{A}{(q+\lambda_1)} + \frac{B}{(q+\lambda_2)}.$$

To solve for A, set $q + \lambda_1 = 0$ in

$$1 = A(q + \lambda_2) + B(q + \lambda_1)$$

This gives

$$A = \frac{1}{q + \lambda_2} = \frac{1}{q + \lambda_1 + \lambda_2 - \lambda_1} = \frac{1}{\lambda_2 - \lambda_1}.$$

Similarly, on setting $q + \lambda_2 = 0$, we find that

$$B = \frac{1}{q + \lambda_1}$$

Thus,

$$\mathbb{E}(e^{-qT}) = \mathbb{E}(e^{-qT_1})\mathbb{E}(e^{-qT_2}) = \frac{\lambda_1 \lambda_2}{(\lambda_2 - \lambda_1)(q + \lambda_1)} + \frac{\lambda_1 \lambda_2}{(\lambda_1 - \lambda_2)(q + \lambda_2)},$$

and since

$$\mathscr{L}^{-1}((q + \lambda_j)^{-1}; t) = e^{-\lambda_j t}$$

It follows that

$$f_T(t) = \lambda_1 \lambda_2 \left(\frac{e^{-\lambda_1 t}}{(\lambda_2 - \lambda_1)} + \frac{e^{-\lambda_2 t}}{(\lambda_1 - \lambda_2)} \right).$$

$\boxed{7, \mathrm{D}}$

2. (a) Data $y_1, \ldots, y_n$ are realisations of random variables $Y_1, \ldots, Y_n$. There is a family of plausible probability distributions for these random variables $\{P_\theta : \theta \in \Theta\}$, called the statistical model. One member of this family, P_{θ^*} say, is the "true" distribution, i.e. the data are one draw from P_{θ^*}, which belongs to the model.

seen ⇓

6, A

(b) For a multiplicative treatment effect on the parameter space $\Gamma = \mathbb{R}^+$, $(g, \gamma) = (\psi, \gamma) \to \gamma\psi \in \Gamma$. The identity element is e such that $(e, \gamma) \to e\gamma = \gamma$, i.e.

seen ⇓

$$(e, \gamma) = (1, \gamma) \to \gamma 1 = \gamma.$$

The group inverse g^{-1} is such that $(g^{-1}, g\gamma) \to g^{-1}g\gamma = \gamma$, i.e.

$$(g^{-1}, g\gamma) \to (\gamma\psi)/\psi = \gamma.$$

The null treatment effect corresponds to the identity element e.

5, C

seen/sim.seen ⇓

(c) Let w_i be the values of a vector of intrinsic variables for the ith pair. Dependence of the distribution of (Y_{i0}, Y_{i1}) on w_i can be incorporated by modelling the pair specific parameter γ_i as a function of w_i. Since $\gamma_i \in \mathbb{R}^+$, a natural modelling assumption is $\gamma_i = \exp(w_i^{\mathrm{T}}\beta)$.

4, A

unseen ⇓

(d) Consider two units u and u' that are comparable in the sense that the values w of their intrinsic variables are the same, i.e. $w(u) = w(u')$. Provided that all relevant aspects have been accounted for in w, it is reasonable to assume that

$$\begin{aligned} \mathsf{pr}_u(A \mid T = 0) &= \mathsf{pr}_{u'}(A \mid T = 0) \\ \mathsf{pr}_u(A \mid T = 1) &= \mathsf{pr}_{u'}(A \mid T = 1). \end{aligned} \tag{2}$$

By conditioning on covariates, there is an implicit comparison of like with like even in the absence of pairing.

5, D

3. (a) (i) From equation (1) and the proportional hazards assumption $h(y;x) = g(x^{\mathrm{T}}\beta)h_0(y)$, the survivor function at y for an individual with covariate x is

seen ⇓

$$\begin{aligned} S(y;x) &= \exp\Big\{-\int_0^y h(u;x)du\Big\} \\ &= \exp\Big\{-g(x^{\mathrm{T}}\beta)\int_0^y h_0(u)du\Big\} \\ &= S_0(y)^{g(x^{\mathrm{T}}\beta)}, \end{aligned}$$

where $S_0(y) = S(y;0)$. From this, the representation $f(y;x) = h(y;x)S(y;x)$, and the proportional hazards assumption,

$$f(y;x) = h_0(y)g(x^{\mathrm{T}}\beta)S_0(y)^{g(x^{\mathrm{T}}\beta)} = f_0(y)g(x^{\mathrm{T}}\beta)S_0(y)^{g(x^{\mathrm{T}}\beta)-1}.$$

4, B

(ii) The hazard functions $h(y;x_i)$ and $h(y;x_j)$ of two individuals with covariate vectors x_i and x_j relate as

seen ⇓

2, A

$$h(y;x_i) = h(y;x_j)\exp\{(x_i - x_j)^{\mathrm{T}}\beta\}. \tag{3}$$

seen/sim.seen ⇓

(iii) From equation (3), the interpretation is that notionally changing the treatment indicator from 0 to 1, keeping all other covariate values equal, multiplies the hazard function by e^{β_p}.

2, A

seen/sim.seen ⇓

(iv) The hazard function for an individual with covariate vector x whose effect is modelled through the rate function as $\rho(x) = \exp(x^{\mathrm{T}}\beta)$ is

$$h(y;x) = f(y;x)/S(y;x) = \alpha\exp(x^{\mathrm{T}}\beta)^{\alpha}y^{\alpha-1}$$

Then the baseline hazard function is $h_0(y) = h(y;0) = \alpha y^{\alpha-1}$ so that

$$\frac{h(y;x)}{h_0(y)} = \exp(\alpha x^{\mathrm{T}}\beta),$$

verifying proportionality of hazards.

4, B

(b) (i) A suitable depiction shows how the timescale for the individual with covariate x relates to that of the individual with covariate at baseline on the scale of the survivor function. E.g. At time $y = 2$ $S(y;x)$ takes the value that $S_0(y)$ took at time 1: $S(2;x) = S_0(1)$, $S(4;x) = S_0(2)$ etc. Visually, $S(y;x)$ is more elongated along the time axis than $S_0(y)$ with the interpretation that the individual with covariate x "wears out" half as fast as the baseline individual.

seen ⇓

4, B

(ii) The survivor function of an individual with covariate x is

seen ⇓

$$S(y;x) = \exp\{-(\exp(x^{\mathrm{T}}\beta)y)^{\alpha}\}$$

and that at baseline is

$$S_0(y) = S(y;0) = \exp(-y^{\alpha}).$$

Thus $S(y;x) = S_0(a(x)y)$ with $a(x) = \exp(x^{\mathrm{T}}\beta)$.

4, A

4. (a) 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

(b) (i) The likelihood function corresponding to n independent observations from an exponential family takes the form unseen ⇓

$$\begin{aligned} L(\alpha,\rho;y) &= \prod_{i=1}^{n} f(y_i;\alpha,\rho) = \prod_{i=1}^{n} \frac{\rho^{\alpha}}{\Gamma(\alpha)} y_i^{\alpha-1} e^{-\rho y_i} \\ &= \frac{\rho^{n\alpha}}{\Gamma(\alpha)^n} \exp\left\{-\rho \sum_{i=1}^{n} y_i + (\alpha-1)\sum_{i=1}^{n} \log y_i\right\}, \end{aligned}$$

showing that $(S,C) = (\sum_{i=1}^{n} \log y_i, \sum_{i=1}^{n} y_i)$ is jointly sufficient. 4, B

(ii) When α is fixed, the likelihood factorizes as unseen ⇓

$$L(\alpha,\rho;y) = \frac{\rho^{n\alpha}}{\Gamma(\alpha)^n} \prod_{i=1}^{n} y_i e^{-\rho c} = h(\alpha;y)g(\alpha,\rho,c),$$

showing that C is sufficient for ρ. 4, C

(ii) Dividing the joint density function of $Y = (Y_1,\ldots,Y_n)$ by the marginal density function for $C = \sum_{i=1}^{n} Y_i$ gives unseen ⇓

$$f_{Y|C}(y|c;\alpha) = \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^n} \prod_{i=1}^{n} (y_i/c)^{\alpha-1}$$

which is free of ρ. A conditional likelihood function for α can be constructed from this conditional density function. 4, D

unseen ⇓

(c) (i) The general form of a partial likelihood for a parameter ψ is the term $L_{\mathsf{pa}}(\psi;y)$ in the factorisation

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

The partial likelihood $L_{\mathsf{pa}}(\psi;y)$ does not depend on λ and is used for inference on ψ, the remainder $L_{\mathsf{r}}(\psi,\lambda;y)$ being discarded. 2, A In the previous question, the full likelihood factors into the contribution from the conditional density function $f_{Y|C}(y|c;\alpha)$ and the contribution from the marginal density function $f_C(c;\alpha,\rho)$. These constitute the partial likelihood and the remainder likelihood respectively.

2, B

(ii) The baseline hazard function h_0, which is an infinite-dimensional unknown parameter, is eliminated through the use of a partial likelihood that depends only on the finite-dimensional parameter vector β. unseen ⇓ 2, A

5. (a) In the absence of covariates, the hazard function $h(y; Z)$ is modelled as a product of a baseline hazard function and a frailty variable Z, which has a distribution on the positive half-line. $\boxed{2, \text{M}}$

(b) Conditional on frailty $Z = z$, the survivor function at y is $S_{Y|Z}(y|z) = \exp\{-zM(y)\}$. The unconditional survivor function is

$$S_Y(y) = \int_0^\infty \exp\{-zM(y)\} f_Z(z) dz$$

which is the Laplace transform $f_Z^*(s) = \mathbb{E}(e^{-sZ})$ of the density function of the frailty variable at $s = M(y)$. $\boxed{2, \text{M}}$

(c) By Bayes's formula, the probability that Z falls in the infinitesimal region $[z, z+dz)$ conditional on survival beyond time y is

$$\text{pr}(z \leq Z \leq z + dz \mid Y \geq y) = \frac{S_{Y|Z}(y \mid z) f_Z(z) dz}{S_Y(y)} = \frac{\exp\{-zM(y)\} f_Z(z) dz}{f_Z^*(M(y))}$$

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

(d) The mean frailty of survivors is typically lower than that of the population overall, because the frail individuals are likely to die early. $\boxed{2, \text{M}}$

(e) A model of the required form would be $h(y; x, z) = zg(x^{\mathrm{T}}\beta)h_0(y)$, so that if the distribution of Z is puts all its mass on a single point, then the resulting constant gets absorbed in the proportionality term $g(x^{\mathrm{T}}\beta)$. $\boxed{2, \text{M}}$

(f) (i) By expressing the survivor in terms of the hazard function,

$$\begin{aligned} S_{Y|Z}(y|z; x) &= \exp\Big(-\int_0^y h_{Y|Z}(y \mid z; x)\Big) \\ &= \exp\Big(-zy - g(x^{\mathrm{T}}\beta)y - \int_0^y h_0(u) du\Big). \end{aligned}$$

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

(ii)

$$\begin{aligned} S_Y(y; x) &= \exp\Big(-g(x^{\mathrm{T}}\beta)y - \int_0^y h_0(u) du\Big) \int_0^\infty \exp(-zy) f_Z(z) dz \\ &= \exp\Big(-g(x^{\mathrm{T}}\beta)y - \int_0^y h_0(u) du\Big) f_Z^*(y). \end{aligned}$$

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

(iii) The Laplace transform of the density function of Z is, by direct calculation

$$f_Z^*(y) = (1 + \alpha y)^{-1/\alpha}$$

Thus

$$S_Y(y; x) = \exp\Big(-g(x^{\mathrm{T}}\beta)y - \int_0^y h_0(u) du\Big)(1 + \alpha y)^{-1/\alpha}$$

and the corresponding hazard function is

$$h_Y(y; x) = -\frac{d}{dy} \log S_Y(y; x) = g(x^{\mathrm{T}}\beta) + h_0(y) + (1 + \alpha y)^{-1}.$$

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

Review of mark distribution:

Total A marks: 32 of 32 marks

Total B marks: 21 of 20 marks

Total C marks: 11 of 12 marks

Total D marks: 16 of 16 marks

Total marks: 100 of 80 marks

Total Mastery marks: 20 of 20 marks

MATH70048 Survival Models Markers Comments

Question 1	Wrong support of integration or computation. An alternative solution was possible for Q1(f), using convolution.
Question 2	Confusing a statistical model with an expression for a regression function, not connecting the intrinsic variables to the proportional hazards model by parameterising the effect (e.g., \gamma_i = \exp(x^t \beta)).
Question 3	Mostly well done. Some students plotted increasing survival functions.
Question 4	Not deriving the expression for the conditional density but explaining with words why it won't depend on \rho. Most students identified the partial likelihood but didn't give motivation and use for simplifying/enabling inference.
Question 5	Not using the definition of conditional density nor mentioning that because frail individuals are more likely to die early, the mean of their distribution will typically be lower. They compared frailty at time zero to frailty at time T, instead.