B.1 Modern Survival Problem sheet 1: Counting processes and martingales

Due at noon, Friday 23 October

(1) We have a point process with intensity

$$\lambda(t) = \begin{cases} 1 & \text{if } 0 \le t < 5, \\ 2 & \text{if } 5 \le t < 10, \\ 0 & \text{otherwise.} \end{cases}$$

Let $(N(t))_{t \ge 0}$ be the counting process and let T_i be the time of the i-th event, or ∞ if there is no i-th event.

(a) i. Find the distribution of the total number of events.

ii. Find the distribution of $T_2 - T_1$.

iii. Show that $T_i \to \infty$ in probability.

(b) In one realisation we observe the points $1.3, 1.7, 3.2, 4.8, 5.5, 6.2, 6.4, 7.0, 7.8, 8.3, 8.4, 8.8, 9.2, 9.9$.

i. What is the compensator $A(t)$ for $N(t)$?

ii. Sketch a typical realisation of $N(t)$.

iii. Sketch the martingale $N(t) - A(t)$.

(2) Let λ be any positive function and $\Lambda(t) = \int_0^t \lambda(s)ds$. Suppose X is a random variable with exponential distribution with parameter 1. Show that $\Lambda^{-1}(X)$ is a random variable with hazard rate $\lambda(t)$.

(3) Suppose X and Y are a pair of random variables with joint density $F(x, y)$. Let $g : \mathbb{R} \to \mathbb{R}$ be a function such that $\mathbb{E}[|g(X)|] < \infty$, and let $\mathcal{Y}$ be the sigma algebra generated by Y. Show that $E[g(X) \mid \mathcal{Y}] = h(Y)$, where

$$h(y) := \frac{\int_{-\infty}^{\infty} g(x) f(x, y) dx}{\int_{-\infty}^{\infty} f(x, y) dx}.$$

Explain the relationship between this formula the conditional expectations you learned in prelims and Part A probability.

(4) Let $(N(t))$ be the counting process associated with the Poisson process with intensity λ, and let $(M(t)) = (N(t) - \lambda t)$ be the associated martingale. Show that λt is the compensator for M^2. (Hint: Expand $M(t)^2 = (N(t) - \lambda t)^2$, and use known properties of the Poisson process.)

(5) Later in the course we will discuss *current status data*, which is a form of extreme censoring. Individuals have an unobserved (assumed i.i.d.) event time U_i. What is observed is a census time C_i (independent of U_i), and $\delta_i = \mathbf{1}_{\{U_i \le C_i\}}$.

(a) Assuming U_i has density f_λ and cdf F_λ, write an expression for the log likelihood of U_i.

(b) Suppose the distribution is exponential, so $f_\lambda(u) = \lambda e^{-\lambda u}$. What is the relative efficiency of estimation (that is, ratio of expected Fisher information) based on the current status data, compared with complete observation of U_i?

(c) Suppose you can choose the distribution of C_i (but still independent of U_i). How would you maximise the expected information?

(6) Suppose we have an inhomogeneous Poisson process $N(t)$, whose intensity starts out as either 1 or 2, each with probability $1/2$, and immediately after each event the intensity is determined again by an independent coin flip. Let $\lambda(t)$ be the intensity at time t.

(a) Suppose $\lambda(t)$ is observed (i.e., $\lambda(t) \in \mathcal{F}_t$). Find the compensator (i.e., the cumulative intensity process) of $N(t)$. Find the predictable variation process and the optional variation process for the martingale $M(t)$ obtained by subtracting the compensator from $N(t)$. Compute $\mathrm{Var}(M(t))$.

(b) Now suppose $\lambda(t)$ is unobserved (i.e., $\lambda(t) \notin \mathcal{F}_t$). Find the compensator. (Remember that the compensator must be $\mathcal{F}_t$-adapted. Find the predictable variation process and the optional variation process of $\widetilde{M}$, where $\widetilde{M}$ is obtained by subtracting the compensator from $N(t)$. Is $\mathrm{Var}(\widetilde{M}(t)) = \mathrm{Var}(M(t))$? Why or why not?