MATH-449 - Biostatistics
EPFL, Spring 2025
Problem Set 11

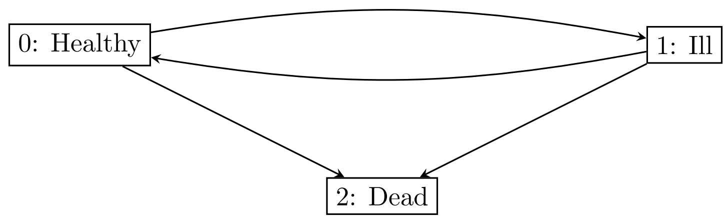

Figure 1: Illness-death model

1. A common example of a multi-state process is an illness-death process. The transition diagram in Figure 1 shows an illness-death process $\{X(t)\}_t$ where people are either healthy (state 0), ill (state 1) or dead (state 2) at any given time t. The arrows indicate the possible transitions: healthy people may become ill and ill people may recover to become healthy as time progresses. All individuals are at risk of dying, i.e., transferring into state 2. Thus, $X(t) \in \{0, 1, 2\}$ for all t.

   (a) We will first consider the time until the *first* event of healthy individuals. We let $\kappa = 2$ signify that the first event is death, and $\kappa = 1$ signify that the first event is illness. We let T^f be the first occurrence of either illness or death and consider an i.i.d. sample

   $$\{(\tilde{T}_i^f, D_i^f, \kappa_i)\}_{i=1}^n$$

   of healthy individuals where

   $$\begin{aligned} D_i^f &= I(T_i^f \leq \tilde{T}_i^f), \\ \tilde{T}_i^f &= \min(T_i^f, T_i^*), \\ N_i^j(t) &= I(\tilde{T}_i^f \leq t, D_i^f = 1, \kappa_i = j), \\ Z_i^f(t) &= I(\tilde{T}_i^f \geq t), \\ \beta^j(t) &= \lim_{h \to 0^+} \frac{1}{h} P(t \leq T^f < t + h, \kappa = j \mid t \leq T^f). \end{aligned}$$

   Let $N^j(t) = \sum_{i=1}^n N_i^j(t)$ and $Z^f(t) = \sum_{i=1}^n Z_i^f(t)$, and suppose that $T^f \perp T^*$ holds, i.e. that a random censoring assumption is satisfied. Answer yes/no on the following claim, and give an explanation for your answer:

   The estimator $\displaystyle\prod_{T_i^f \leq t} \left\{1 - \frac{\Delta N^2(T_i^f)}{Z^f(T_i^f)}\right\}$ is an estimator of the survival function at time t.

   (b) We will now return to the illness-death process X shown in the diagram in Figure 1. In the lectures we defined the transition intensities $\alpha_{ij}(t)$ by

   $$\alpha_{ij}(t) = \lim_{h \to 0^+} \frac{1}{h} P_{ij}(t, t + h) = \lim_{h \to 0^+} \frac{1}{h} P(X(t + h) = j \mid X(t) = i)$$

   for transitioning from state i to state j when $i \neq j$, and $\alpha_{ii}(t) = -\sum_{j \neq i} \alpha_{ij}(t)$. Use the definition of the transition intensity matrix,

   $$\boldsymbol{\alpha}(t) = \lim_{h \to 0^+} \frac{1}{h}(\mathbf{P}(t, t + h) - \mathbf{I}),$$

   to show that the (i, j)'th element of $\boldsymbol{\alpha}(t)$ is $\alpha_{ij}(t)$, i.e. that $\big(\boldsymbol{\alpha}(t)\big)_{ij} = \alpha_{ij}(t)$.

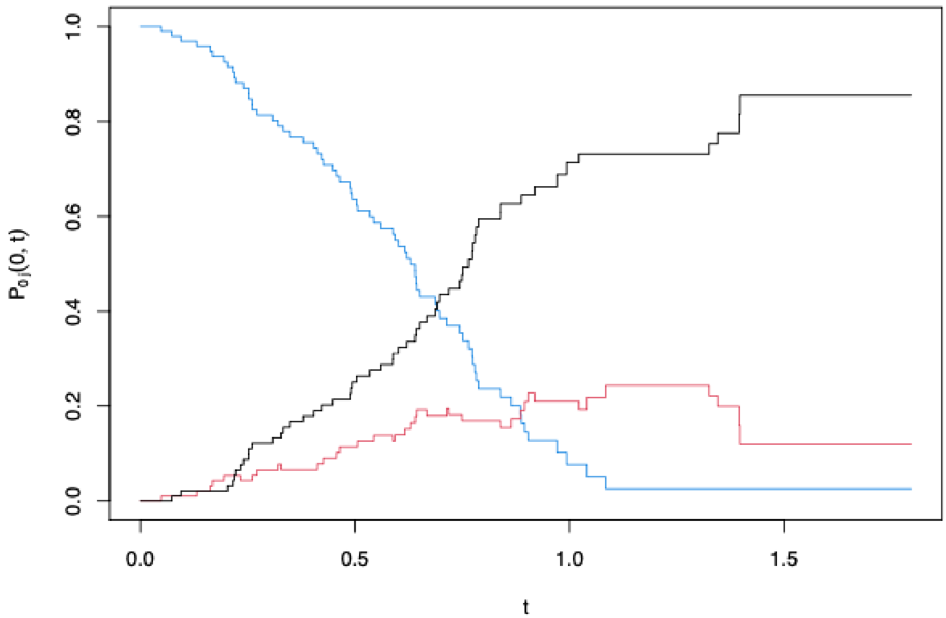

Figure 2: Plot of estimated transition probabilities based on a simulated illness death model.

(c) Is $\alpha_{ij}(t)$ zero for any (i, j) in the example shown in the graph in Figure 1? Write out the matrix $\boldsymbol{\alpha}(t)$. Use the Kolmogorov forward equation to express $(P_{00}(s,t), P_{01}(s,t))^\top$ as an integral equation on the form

$$\begin{pmatrix} P_{00}(s,t) \\ P_{01}(s,t) \end{pmatrix} = \begin{pmatrix} f^0(s,t) \\ f^1(s,t) \end{pmatrix} + \int_s^t G(t,u) \begin{pmatrix} P_{00}(s,u) \\ P_{01}(s,u) \end{pmatrix} du$$

where G is a 2×2 matrix. What is f^0, f^1 and G?

(d) Attached is a simulation of an illness-death process where all individuals started in state 0 (healthy). Estimates of the transition probabilities $P_{0j}(0,t)$ for $j \in \{0, 1, 2\}$, obtained from the Aalen-Johansen estimator, are plotted over time t in Figure 2. State which of the curves (blue, black, red) correspond to which of the quantities P_{00}, P_{01} and P_{02}. Explain your reasoning.

(e) The transition intensity matrix $\boldsymbol{\alpha}$ used for generating Figure 2 contains one zero more than the transition intensity matrix you wrote down in point (c). That is, there is a combination (i, j) such that $\alpha_{ij} = 0$ in the data generating model that gave rise to Figure 2, but the diagram in Figure 1 suggests that particular α_{ij} is not zero. State which combination (i, j) it is, and explain why Figure 2 indicates that the transition intensity $\alpha_{ij}(t)$ is zero.

2. Compute $P(Z = 0|X_0 = 10, Y_0 = 1)$, $P(Z = 1|X_0 = 10, Y_0 = 1)$, $P(Z = 2|X_0 = 10, Y_0 = 1)$ and $P(Z = 3|X_0 = 10, Y_0 = 1)$ for the Reed-Frost model.

Hint: Compute the probability for each epidemic chain giving $Z = k$ separately and then add these probabilities.

3. Show that $y(t) \to 0$ as $t \to \infty$ for the model of Kermack and McKendrick given by the differential equations

$$\begin{aligned} x'(t) &= -\lambda x(t)y(t) \\ y'(t) &= \lambda x(t)y(t) - \gamma y(t) \\ z'(t) &= \gamma y(t), \end{aligned}$$

for some $\lambda, \gamma > 0$