MATH-449 - Biostatistics
EPFL, Spring 2023
Problem Set 4 - Answer Key

1. (Exercise 2.4 in ABG 2008) Let M be a discrete time martingale with respect to the filtration $\mathcal{F}_n$, for $n \in \{0, 1, 2, \cdots\}$, and suppose $M_0 = 0$. Prove that $M^2 - \langle M \rangle$ is a martingale with respect to the filtration $\mathcal{F}$, that is, that $E(M_n^2 - \langle M \rangle_n | \mathcal{F}_{n-1}) = M_{n-1}^2 - \langle M \rangle_{n-1}$.

   Solution We have $\langle M \rangle_n = \sum_{i=1}^n E[(M_i - M_{i-1})^2 | \mathcal{F}_{i-1}]$ and $\langle M \rangle_0 = 0$ by definition, which means that $M_0^2 - \langle M \rangle_0 = 0$. For $n > 0$,

$$
\begin{aligned}
E\big[M_n^2 - \langle M \rangle_n | \mathcal{F}_{n-1}\big] &= E\Big[(M_n - M_{n-1} + M_{n-1})^2 - E[(M_n - M_{n-1})^2 | \mathcal{F}_{n-1}] - \langle M \rangle_{n-1} \Big| \mathcal{F}_{n-1}\Big] \\
&= E\Big[(M_n - M_{n-1})^2 + M_{n-1}^2 - E[(M_n - M_{n-1})^2 | \mathcal{F}_{n-1}] \Big| \mathcal{F}_{n-1}\Big] - \langle M \rangle_{n-1} \\
&= E\Big[(M_n - M_{n-1})^2 - E[(M_n - M_{n-1})^2 | \mathcal{F}_{n-1}] \Big| \mathcal{F}_{n-1}\Big] + M_{n-1}^2 - \langle M \rangle_{n-1} \\
&= M_{n-1}^2 - \langle M \rangle_{n-1}
\end{aligned}
$$

   where we used that $E[(M_n - M_{n-1})M_{n-1}] = 0$, and that $\langle M \rangle_{n-1}$ and M_{n-1}^2 are $\mathcal{F}_{n-1}$-measurable.

2. Suppose we have n independent survival times $\{T_i\}_{i=1}^n$, where T_i corresponds to the time of death of individual i. Suppose we somehow could observe each individual from $t = 0$ up to his/her time of death.

   In the lectures you learned that a counting process $\{N(t)\}_{t \geq 0}$ is an increasing right-continuous integer-valued stochastic process such that $N(0) = 0$. Write down the counting process N_i^c (that "counts" the death of individual i) in terms of T_i.

   Solution N_i^c has at most one jump, at time T_i. As $N_i^c(t)$ is right-continuous it must take the form $N_i^c(t) = I(T_i \leq t)$

   You have also learned about the *intensity process* λ of a counting process N with respect to a filtration $\mathcal{F}$. It is informally defined through the relationship $\lambda(t)dt = E[dN(t)|\mathcal{F}_t]$.

3. Argue that the intensity of N^c is given by $\lambda_i^c(t) = \alpha_i(t) Z_i(t)$, where $Z_i(t) = I(T_i \geq t)$.

   In general, if the intensity $\lambda(t)$ of a counting process $N(t)$ with respect to $\mathcal{F}_t$ can be written on the form

$$
\lambda(t) = \alpha(t) \cdot Z(t),
$$

   where α is an unknown deterministic function and Z is an $\mathcal{F}_t$-predictable[§] function that does not depend on α, $N(t)$ is said to satisfy the *multiplicative intensity model*[*].

4. (Exercise 1.10 in ABG 2008) Consider the scenario in Exercise 2, and let $\mathcal{F}_t^c$ be the filtration generated by $\{N_i^c(s), s \leq t, i = 1, \cdots, n\}$. In the lectures we have seen that the intensity of N_i^c with respect to $\mathcal{F}^c$ in this case is $\lambda_i^c(t) = E[dN_i^c(t)|\mathcal{F}_t^c] = \alpha_i(t) Z_i(t)$, where $\alpha_i(t)$ is the hazard function of individual i and $Z_i(t) = I(T_i \geq t)$. Consider the aggregated counting process $N^c(t) = \sum_{i=1}^n N_i^c(t)$.

   i) Let $\{\eta_i(t)\}_{i=1}^n$ be known, positive, continuous functions. Find the intensity process of N^c with respect to $\mathcal{F}_t^c$ when α_i take the following forms:

      a) $\alpha_i(t) = \alpha(t)$
      b) $\alpha_i(t) = \eta_i(t)\alpha(t)$
      c) $\alpha_i(t) = \alpha(t) + \eta_i(t)$

[§] Recall that this holds when Z is left-continuous and adapted to $\mathcal{F}$, i.e. that all the information needed to know the value of Z at time t is contained in $\mathcal{F}_t$.

[*] We will later derive estimators for the unknown function α under the multiplicative intensity model.

ii) For which of the three cases in i) does N^c satisfy the multiplicative intensity model?

Solution We have

$$\begin{aligned}\lambda^c(t)dt &= E[dN^c(t)|\mathcal{F}_t^c] \\ &= \sum_{i=1}^n E[dN_i^c(t)|\mathcal{F}_t^c] \\ &= \sum_{i=1}^n \alpha_i(t)Z_i(t)dt.\end{aligned}$$

Define $Z(t) = \sum_{i=1}^n Z_i(t)$ and $Z_\eta(t) = \sum_{i=1}^n \eta_i(t)Z_i(t)$. Note first that Z_i is left-continuous. Also, all the information needed to determine whether T_i have happened at t is contained in $\mathcal{F}_t^c$. We thus have

i) a) $\lambda^c(t) = \alpha(t)Z(t)$
   b) $\lambda^c(t) = \alpha(t)Z_\eta(t)$
   c) $\lambda^c(t) = \alpha(t)Z(t) + Z_\eta(t)$.

ii) a) Yes
   b) Yes
   c) No. For any representation $\lambda^c(t) = \tilde{\alpha}(t)\tilde{Z}(t)$, either $\tilde{\alpha}$ will not be deterministic or $\tilde{Z}$ will be a function of α.

5. Let N be a nonhomogeneous Poisson process with deterministic intensity function $\alpha(t)$. Define $H(t) = \int_0^t \alpha(s)ds$. The following two points i)-ii) provide equivalent definitions of such a process:

   i)
   - $N(t) - N(s) \sim \text{Poisson}(H(t) - H(s))$ for $s < t$
   - $N(t) - N(s)$ is independent of $\mathcal{F}_s$ for $s < t$

   ii)
   $$\begin{aligned}P(N_{t+\delta} - N_t = 1|\mathcal{F}_t) &= \alpha(t)\delta + o\big(\delta^2\big) \\ P(N_{t+\delta} - N_t = 0|\mathcal{F}_t) &= 1 - \alpha(t)\delta + o\big(\delta^2\big)\end{aligned}$$
   as $\delta \longrightarrow 0^+$.

   Here, $\mathcal{F}$ is the filtration generated by N. The second condition in i) implies that $E[N(t) - N(s)|\mathcal{F}_s] = E[N(t) - N(s)]$.

   a) Show that $M = N - H$ is a martingale with respect to $\mathcal{F}$.*
   b) Show that the increments of M are uncorrelated, i.e. that, for $v \le u \le s \le t$, †
   $$E[(M(t) - M(s))(M(u) - M(v))] = 0.$$

   Suppose that N is only recorded up to the deterministic time X, and define $N^\star(t) = N(\min\{t, X\})$. Thus, $N^\star$ is the process N censored at X.

   c) Argue that $N^\star(t)$ is the observed number of jumps of N up to time t, and demonstrate that $N^\star$ satisfies the multiplicative intensity model.‡
   d) Suppose now that X is a random variable. Verify that the conclusion in c) holds when $\{X \le t\} \in \mathcal{F}_t$ for each t, or equivalently, that $I(X \le \cdot)$ is adapted to $\mathcal{F}$.§

*Hint: A Poisson distributed variable with parameter $\lambda > 0$ has mean λ.
†Note: this is true for any martingale M, not just the one from a).
‡Hint: start with definition ii). Alternatively, you may find it helpful to use $N^\star(t) = \int_0^t I(X \ge s)dN_s$.
§X is then called a *stopping time* with respect to $\mathcal{F}$. Heuristically, $\mathcal{F}_t$ contains enough information to determine whether X has occurred by t.

Solution

a) Since $N(t)-N(s)$ is Poisson distributed with parameter $H(t)-H(s)$ we have $E[N(t)-N(s)] = H(t)-H(s)$. Moreover, as $N(t)-N(s)$ is independent of $\mathcal{F}_s$ we have that $E[N(t)-N(s)|\mathcal{F}_s] = E[N(t)-N(s)]$ Combining those two observations we get

$$\begin{aligned} E\big[N(t)-H(t)-\big(N(s)-H(s)\big)|\mathcal{F}_s\big] &= E\big[N(t)-N(s)|\mathcal{F}_s\big]-(H(t)-H(s)) \\ &= E\big[N(t)-N(s)\big]-(H(t)-H(s)) \\ &= 0, \end{aligned}$$

where we used the fact that $H(t)$ and $H(s)$ are $\mathcal{F}$-measurable (since they are deterministic functions) in the first line, that $N(t)-N(s)$ is independent of $\mathcal{F}_s$ in the second line, and the expectation of a Poisson distributed variable in the third line.

b) We have

$$\begin{aligned} E\big[(M(t)-M(s))(M(u)-M(v))\big] &= E\Big[E\big[(M(t)-M(s))(M(u)-M(v))|\mathcal{F}_s\big]\Big] \\ &= E\Big[E\big[(M(t)-M(s))|\mathcal{F}_s\big](M(u)-M(v))\Big] \\ &= E\Big[\big(E[M(t)|\mathcal{F}_s]-M(s)\big)(M(u)-M(v))\Big] \\ &= 0, \end{aligned}$$

where we used the law of total expectation in the first line, the fact that $M(u)-M(v)$ is $\mathcal{F}_s$-measurable in the second line (since $u, v \leq s$), and the definition of a martingale in the last line.

c) Definition ii) leads to the infinitesimal result

$$E[dN(t)|\mathcal{F}_t] = \alpha(t)dt.$$

Clearly, the intensity of $N^\star$ coincides with the intensity of N for $X \geq t$, while it is zero for $X < t$. Thus

$$E[dN^\star(t)|\mathcal{F}_t] = \alpha(t)I(X \geq t)dt,$$

which means that the intensity of $N^\star(t)$ with respect to $\mathcal{F}_t$ is $\alpha(t)I(X \geq t)$, and $N^\star$ satisfies the multiplicative intensity model. Alternatively, use the identity $N^\star(t) = \int_0^t I(X \geq s)dN(s)$ directly. This leads to

$$\begin{aligned} E[dN^\star(t)|\mathcal{F}_t] &= E[I(X \geq t)dN(t)|\mathcal{F}_t] \\ &= I(X \geq t)E[dN(t)|\mathcal{F}_t] \\ &= I(X \geq t)\alpha(t)dt, \end{aligned}$$

where we used the fact that $I(X \geq \cdot)$ is adapted (since it is deterministic) in the second line.

d) We still have that the intensity of $N^\star(t)$ is $\lambda^\star(t) = \alpha(t)I(X \geq t)$. By this choice of X, $I(X \geq \cdot)$ is both adapted to $\mathcal{F}$ and left-continuous, and $\lambda^\star(t)$ satisfies the multiplicative intensity model.

Solution

a) Since $N(t)-N(s)$ is Poisson distributed with parameter $H(t)-H(s)$ we have $E[N(t)-N(s)] = H(t)-H(s)$. Moreover, as $N(t)-N(s)$ is independent of $\mathcal{F}_s$ we have that $E[N(t)-N(s)|\mathcal{F}_s] = E[N(t)-N(s)]$ Combining those two observations we get

$$\begin{aligned} E\big[N(t)-H(t)-\big(N(s)-H(s)\big)|\mathcal{F}_s\big] &= E\big[N(t)-N(s)|\mathcal{F}_s\big]-(H(t)-H(s)) \\ &= E\big[N(t)-N(s)\big]-(H(t)-H(s)) \\ &= 0, \end{aligned}$$

where we used the fact that $H(t)$ and $H(s)$ are $\mathcal{F}$-measurable (since they are deterministic functions) in the first line, that $N(t) - N(s)$ is independent of $\mathcal{F}_s$ in the second line, and the expectation of a Poisson distributed variable in the third line.

b) We have

$$\begin{aligned} E\big[(M(t) - M(s))(M(u) - M(v))\big] &= E\Big[E\big[(M(t) - M(s))(M(u) - M(v))|\mathcal{F}_s\big]\Big] \\ &= E\Big[E\big[(M(t) - M(s))|\mathcal{F}_s\big](M(u) - M(v))\Big] \\ &= E\Big[\big(E[M(t)|\mathcal{F}_s] - M(s)\big)(M(u) - M(v))\Big] \\ &= 0, \end{aligned}$$

where we used the law of total expectation in the first line, the fact that $M(u) - M(v)$ is $\mathcal{F}_s$-measurable in the second line (since $u, v \leq s$), and the definition of a martingale in the last line.

c) Definition ii) leads to the infinitesimal result

$$E[dN(t)|\mathcal{F}_t] = \alpha(t)dt.$$

Clearly, the intensity of $N^\star$ coincides with the intensity of N for $X \geq t$, while it is zero for $X < t$. Thus

$$E[dN^\star(t)|\mathcal{F}_t] = \alpha(t)I(X \geq t)dt,$$

which means that the intensity of $N^\star(t)$ with respect to $\mathcal{F}_t$ is $\alpha(t)I(X \geq t)$, and $N^\star$ satisfies the multiplicative intensity model. Alternatively, use the identity $N^\star(t) = \int_0^t I(X \geq s)dN(s)$ directly. This leads to

$$\begin{aligned} E[dN^\star(t)|\mathcal{F}_t] &= E[I(X \geq t)dN(t)|\mathcal{F}_t] \\ &= I(X \geq t)E[dN(t)|\mathcal{F}_t] \\ &= I(X \geq t)\alpha(t)dt, \end{aligned}$$

where we used the fact that $I(X \geq \cdot)$ is adapted (since it is deterministic) in the second line.

d) We still have that the intensity of $N^\star(t)$ is $\lambda^\star(t) = \alpha(t)I(X \geq t)$. By this choice of X, $I(X \geq \cdot)$ is both adapted to $\mathcal{F}$ and left-continuous, and $\lambda^\star(t)$ satisfies the multiplicative intensity model.

6. In this problem we will use the definition of the optional variation process $[\cdot]$ from the lecture notes. Thus, we will need to take limits $[G](t) = \lim_{n\to\infty} \sum_{k=1}^n \big(G(kt/n) - G((k-1)t/n)\big)^2$ (in probability) of processes G.

Let $\{N(t) : t \in [0, \tau]\}$ be a counting process. Let λ be the intensity of N with respect to some filtration $\mathcal{F}$, so that $\Lambda(t) = \int_0^t \lambda(s)ds$ is the cumulative intensity, and $M = N - \Lambda$ is a martingale with respect to $\mathcal{F}$. Assume that $\int_0^\tau \lambda(s)^2 ds \leq K$ for some constant K.

a) Show that N has a finite number of jumps with probability 1. Hint: start by looking at $E[N(\tau)]$, use that M is a martingale and that $\int_0^\tau \lambda(s)^2 ds \leq K$.[¶]

b) Show that the optional variation process $[N]$ is equal to N (recall that there are no tied event times, so that $N(t) - N(t-) \leq 1$ for all t).

Solution

[¶]Hint: Use also the inequality $\big(\int_a^b f(s)ds\big)^2 \leq (b-a)\int_a^b f(s)^2 ds$.

a) We have that $\int_0^\tau \lambda(s)ds \leq \sqrt{\int_0^\tau 1^2 ds \cdot \int_0^\tau \lambda(s)^2 ds} \leq \sqrt{\tau K}$, where we used the Cauchy-Schwarz inequality, and the assumption $\int_0^\tau \lambda(s)^2 ds \leq K$ for some constant K. We can thus conclude that $E[N(\tau)] = E[\int_0^\tau \lambda(s)ds] \leq \sqrt{\tau K}$. Now, this implies that $P(N(\tau) < \infty) = 1$ (otherwise

$$E[N(\tau)]$$

wouldn't be finite). Thus, with probability 1, a realisation of N will have a finite number of jumps.

b) Choose n' large enough so that each of the intervals $\{[(k-1)t/n', kt/n')\}_{k=1}^{n'}$ has at most one event (we have seen that such an n' is guaranteed to exist). Thus,

$$\big(N(kt/n') - N((k-1)t/n')\big)^2 = \begin{cases} 1, & \text{if there is an event in} \big[(k-1)t/n', kt/n'\big) \\ 0, & \text{otherwise.} \end{cases}$$

Since N has a total of $N(t)$ jumps of size 1 up to t, the sum reduces to

$$\sum_{k=1}^{n} \big(N(kt/n) - N((k-1)t/n)\big)^2 = N(t),$$

whenever $n \geq n'$. Hence,

$$[N](t) = \lim_{n\to\infty} \sum_{k=1}^{n} \big(N(kt/n) - N((k-1)t/n)\big)^2 = N(t).$$

7. Suppose $M = \{M_0, M_1, M_2, \dots\}$ is a discrete Martingale. Show that $Cov(M_m, M_n - M_m) = 0, \forall n > m$.

Solution