

## MAST30001 Stochastic Modelling

### Tutorial Sheet 11

1. The lifetime in hours of a saw blade in a lumber mill is a random variable with density proportional to  $x(2-x)$  on  $0 < x < 2$ . When the mill opened a new blade was put in the saw. Every time a blade fails it is immediately replaced with a new one. Since the mill is automated, it operates 24 hours a day, seven days a week.
  - (a) Model the number  $N_t$  of saw blades that have been replaced  $t$  hours after the mill opened as a renewal process and determine the density, mean, and variance of the random times between renewals.
  - (b) On average, about how many replacement blades for the saw will the mill go through in the first 30 days of opening?
  - (c) Give an interval around your estimate from (b) that will have a 95% chance of covering the true number of saws needed in the first 30 days.
  - (d) If you show up at the mill at the end of day 30, about what is the mean and variance of the amount of time you'll have to wait for the current blade in use to be replaced?

**Ans.**

- (a) If  $X_1, \dots$  are i.i.d. with distribution the lifetime of the saw then

$$N_t = \max \left\{ n : \sum_{i=1}^n X_i \leq t \right\}$$

defines a renewal process with inter-arrival times  $X_i$ . The density of  $X_i$  is

$$\frac{3}{4}x(2-x) \quad \text{for } 0 < x < 2,$$

with mean  $\mu = 1$  and variance  $\sigma^2 = 1/5$ .

- (b) Renewal theory says that for large  $t$ ,  $N_t/t \approx 1/\mu = 1$ ; here  $t = 30 * 24 = 720$  so we expect to need about that many saws up through hour 720.

- (c) The fluctuations around the mean are given by the renewal CLT which states that for large  $t$

$$\frac{N_t - t/\mu}{\sqrt{t\sigma^2/\mu^3}}$$

is roughly standard normal. So the fluctuations around our estimate for  $t = 720$  and the parameters above are roughly normal with SD=12 and so the 95% CI is

$$720 \pm (1.96) * 12.$$

- (d) Defining  $T_k = \sum_{i=1}^k X_i$ , we know that for large  $t$  and  $Y_t := T_{N_t+1} - t$  roughly has density  $(1 - F(x))/\mu$  for  $0 < x < 2$ . In this case this density on this support is

$$1 - 3x^2/4 + x^3/4,$$

and now its only a matter of computing the first and second moments to find the mean is  $3/5$  and the variance is  $13/75$ .

2. A faculty course advisor is giving advice to students doing either course A or course B. Course A is rather structured and so advice is simple. Course B has more flexibility and so students need more time to select subjects. Moreover, in some complicated situations advising time can be more than 15 minutes and for these cases, the student is sent to talk to a faculty officer after 15 minutes. We assume that the distributions of the time (in hours) that the advisor spends talking to a given student have densities

$$\begin{aligned} \text{(A)} \quad f_A(x) &= 32(1/4 - x), \quad x \in (0, 1/4), \\ \text{(B)} \quad f_B(x) &= 32x, \quad x \in (0, 1/4). \end{aligned}$$

On each day, either only the course A or course B students can come to the course advice session. We assume advisors work without breaks, there are always students waiting for advice, and that the time that advice sessions last are i.i.d. Denote by  $N_t$  the number of students an advisor has finished talking to  $t$  hours into the day. Answer the following questions for both cases (A) and (B).

- (a) About how many students can be advised in a typical 8 hour work day?
- (b) Determine an interval that contains  $N_8$  with roughly 95% probability.
- (c) What is the density of the limiting (as  $t \rightarrow \infty$ ) residual lifetime  $T_{N_t+1} - t$ ?
- (d) What is the joint density of the limiting (as  $t \rightarrow \infty$ ) residual lifetime  $T_{N_t+1} - t$  and age  $t - T_{N_t}$ ?

**Ans.** The analysis is similar for (A) and (B) so we only do (A) (also note this is question 2 in Chapter 8 of Borovkov and the solution there works out (B)).

(a)  $(N_t)_{t \geq 0}$  is a renewal process having interarrival variables  $\tau_i$  having density  $f_A$ . Thus the renewal LLN says  $E[\tau_1] = 1/12$ . Since  $N_t/t \rightarrow 1/E[\tau_1]$ , we estimate  $N_8 \approx 8 \times 12 = 96$ .

(b) The renewal CLT says  $N_8$  is roughly normal with mean 96 and variance

$$\frac{8 \times \text{Var}(\tau_1)}{E[\tau_1]^3} = 48,$$

since  $\text{Var}(\tau_1) = 1/288$ . Thus we expect  $N_8$  to be in the interval

$$96 \pm (1.96) \times \sqrt{48}$$

with probability 0.95.

(c) From lecture, the limiting density is given by

$$g(x) = \frac{1}{E[\tau_1]} P(\tau_1 > x) = 12(1 - 4x)^2, \quad 0 < x < 1/4.$$

(d) Again from lecture, the joint density is

$$h(x, y) = \frac{f_A(x+y)}{\mu} = 12 \times 32(1/4 - x - y), \quad x, y > 0, \quad 0 < x + y < 1/4.$$

3. (Thinned Renewal Process) Consider a renewal process with mean interarrival time  $\mu$ . Suppose that each event of this process is independently “counted” with probability  $p$ . Let  $N_t$  denote the number of counted events by time  $t$ ;  $t > 0$ .

- (a) Is  $(N_t)_{t \geq 0}$  a renewal process?  
 (b) What is  $\lim_{t \rightarrow \infty} N_t/t$ ?

**Ans.**

(a) The process is a renewal process with interarrival times distributed as  $\sum_{i=1}^M \tau_i$ , where  $\tau_i$  is the original renewal process interarrival distribution and  $M$  is a geometric distribution with parameter  $p$  started from one and independent of the  $\tau_i$ .

(b) Renewal results from lecture imply  $N_t/t \rightarrow 1/E[\sum_{i=1}^M \tau_i]$  and

$$E \left[ \sum_{i=1}^M \tau_i \right] = E[M]E[\tau_1] = \mu/p.$$

That is,  $\lim_{t \rightarrow \infty} N_t/t = p/\mu$ .

4. (Delayed renewal process) A renewal process for which the time until the initial renewal, say  $\tau_0$ , has a different distribution than the remaining interarrival times, say  $\tau_1, \tau_2, \dots$ , is called a delayed (or a general) renewal process. If  $N_t$  is the number of renewals of a delayed renewal process in  $[0, t]$ , then show that  $\lim_{t \rightarrow \infty} N_t/t \rightarrow 1/\mu$ , where  $\mu = E[\tau_1]$ .

**Ans.** If  $\tilde{N}_t = N_{t+\tau_0} - 1$ , then  $(\tilde{N}_t)_{t \geq 0}$  is a standard renewal process with no delay and so we know that  $\tilde{N}_t/t \rightarrow 1/\mu$ . Now, for  $t > \tau_0$ ,

$$N_t = \tilde{N}_{t-\tau_0} + 1$$

and so

$$\frac{N_t}{t} = \frac{\tilde{N}_{t-\tau_0}}{t-\tau_0} \frac{t-\tau_0}{t} + \frac{1}{t}.$$

As  $t \rightarrow \infty$ ,

$$\frac{\tilde{N}_{t-\tau_0}}{t-\tau_0} \rightarrow \frac{1}{\mu}, \quad \frac{t-\tau_0}{t} \rightarrow 1, \quad \frac{1}{t} \rightarrow 0,$$

and so  $N_t/t \rightarrow 1/\mu$ .

5. The phenomenon in a renewal process  $(N_t)_{t \geq 0}$ , that  $T_{N_t+1} - T_{N_t}$ , the length of a typical interval straddling a large value of  $t$ , is larger than the typical interval length,  $\tau_i$ , is called *size biasing*. Intervals aren't selected at random to straddle  $t$ , but longer intervals are more likely. Size biasing arises frequently in statistical sampling. Consider a remote tourist destination that wants to estimate the average length of a tourist's visit. Two sampling schemes are up for discussion. The first scheme uses the estimate of the average of the lengths of tourists' visit from a randomly chosen sample of tourists leaving the airport of the destination while the second scheme uses the estimate of the average of the lengths of tourists' visit from a randomly chosen sample of tourists staying at hotels. These averages will not be the same. Why? Which is bigger? Which scheme is the most appropriate to estimate the average length of a tourist's visit?

**Ans.** The average sampled from the hotels is larger since it's more likely to capture people with longer stays; that is, the chance that a given person is sampled in a hotel is proportional to the length of their stay in that hotel. Sampling from the airport gives a better estimate since each person flies out of the airport exactly once and the length of time spent in the airport doesn't vary greatly from person to person (or at least not as much as hotel stays).