EE5137 : Stochastic Processes (Spring 2022)
Some Notes on the Distribution of S_n

Vincent Y. F. Tan

February 11, 2022

In this document, we provide a more detailed explanation of (i) the fact that the only random variables on $[0,\infty)$ that are memoryless are $\mathrm{Exp}(\lambda)$ (mentioned this in Lecture 4) and (ii) how to derive the distribution of S_n, the n^{th} arrival epoch of a Poisson process (to be discussed in Lecture 5).

1 Memorylessness

A random variable X is *memoryless* if X is a positive random variable and for all $x, t > 0$, we have

$$\Pr(X > x + t) = \Pr(X > x)\Pr(X > t). \tag{1}$$

It is easy to check that for $X \sim \mathrm{Exp}(\lambda)$ for any $\lambda > 0$ is memoryless. We want to show the converse, i.e., if X is supported on $[0,\infty)$, the only random variables that are memoryless are $\mathrm{Exp}(\lambda)$ for some λ. Notice from (1) and the definition $h(x) = \log \Pr(X > x)$, we have that

$$h(x + y) = h(x) + h(y) \tag{2}$$

and furthermore, $h : [0,\infty) \to \mathbb{R}$ is non-increasing (since the CDF is non-decreasing). We aim to show here that the only functions that satisfy (2) and are non-increasing are linear, i.e., $h(x) = cx$ for some $c \in \mathbb{R}$. We will show this in a sequence of steps. We assume that $h(1) = c$ for some $c \in \mathbb{R}$; this is for normalization.

First, we show that for $n \in \mathbb{N}$, $h(n) = nc$. We use induction. We know $h(1) = c$ and the inductive hypothesis is that $h(n) = nc$, which is satisfied for $n = 1$. We then have $h(n + 1) = h(n) + h(1) = nc + c = (n+1)c$. Thus if the hypothesis is satisfied for n it is also satisfied for $n+1$, which verifies that it is satisfied for all positive integer n.

Next we show that any $j \in \mathbb{N}$, $h(1/j) = c/j$. Repeatedly adding $h(1/j)$ to itself, we get $h(2/j) = h(1/j) + h(1/j) = 2h(1/j)$, $h(3/j) = h(2/j) + h(1/j) = 3h(1/j)$ and so forth to $h(1) = h(j/j) = jh(1/j)$. Thus $h(1/j) = c/j$.

Next, we show that for any $k, j \in \mathbb{N}$, $h(k/j) = ck/j$. Since $h(1/j) = c/j$, for each positive integer j, we can use induction on positive integers k for any given $j > 0$ to get $h(k/j) = ck/j$.

So we have shown that $h(x) = cx$ for all positive rationals $x \in \mathbb{Q}$. We now extend the claim to all positive reals. Let $x > 0$ be a real number and let $\{x_i\}$ be a sequence of increasing rational numbers approaching x, i.e., $x_i \uparrow x$. Thus, $h(x_i) \geq h(x)$ for all i (because h is non-increasing). Then $\liminf_{i\to\infty} h(x_i) = c \liminf_{i\to\infty} x_i = cx$. Thus $\liminf_{i\to\infty} h(x_i) \geq cx$. If we look at a similarly decreasing sequence $\{x_i'\}$, we see that $\limsup_{i\to\infty} h(x_i') \leq cx$. Since $\limsup_{i\to\infty} h(x_i') \leq h(x) \leq \liminf_{i\to\infty} h(x_i)$ and the liminf and limsup are both equal to cx, this proves that h is a linear function. *This analysis will not be tested.*

2 Distribution of S_n

Recall that $S_n = \sum_{i=1}^n X_i$ where $\{X_i\}_{i=1}^\infty$ is a sequence of i.i.d. interarrival times, each distributed as $\mathrm{Exp}(\lambda)$, i.e., the exponential distribution with rate λ or

$$f_{X_i}(x) = \lambda \exp(-\lambda x), \qquad x \geq 0, \tag{3}$$

for all $i \in \mathbb{N}$.

Of course, the most straightforward way to derive f_{S_n} is to convolve $\mathrm{Exp}(\lambda)$ a total of n times. But this is too cumbersome. The book suggests the following procedure.

First consider the joint distribution of (X_1, S_2), which is the same as (S_1, S_2). This is given by

$$f_{X_1,S_2}(x_1,s_2) = f_{X_1}(x_1) f_{S_2|X_1}(s_2|x_1) = \lambda \exp(-\lambda x_1) \cdot \lambda \exp(-\lambda(s_2 - x_1)), \qquad 0 \le x_1 \le s_2. \tag{4}$$

This is of course equivalent to

$$f_{X_1,S_2}(x_1,s_2) = \lambda^2 \exp(-\lambda s_2) \qquad 0 \le x_1 \le s_2. \tag{5}$$

Hence, given we know $S_2 = s_2$, the distribution of X_1 is uniform over $[0, s_2]$. Clearly, since $X_1 = S_1$, we also have

$$f_{S_1,S_2}(s_1,s_2) = \lambda^2 \exp(-\lambda s_2) \qquad 0 \le s_1 \le s_2. \tag{6}$$

To obtain $f_{S_2}(s_2)$, we simply integrate out s_1, i.e.,

$$f_{S_2}(s_2) = \int_0^{s_2} \lambda^2 \exp(-\lambda s_2)\, \mathrm{d}s_1 = \lambda^2 s_2 \exp(-\lambda s_2), \qquad s_2 \in [0,\infty) \tag{7}$$

This is the *Erlang density of order* 2.

Now we claim that in the general case

$$f_{S_n}(s_n) = \frac{\lambda^n s_n^{n-1} \exp(-n s_n)}{(n-1)!}, \qquad s_n \in [0,\infty). \tag{8}$$

This is the *Erlang density of order* n. To show this, consider the joint distribution of $S_1, S_2, \ldots S_n$. We claim that it is

$$f_{S_1,\ldots,S_n}(s_1,\ldots,s_n) = \lambda^n \exp(-\lambda s_n), \qquad 0 \le s_1 \le s_2 \le \ldots \le s_n. \tag{9}$$

This clearly checks out for $n = 2$ by the derivation leading to (6). The general case proceeds by induction. We assume that (9) is true for the n^{th} proposition (the induction hypothesis). Then consider

$$\begin{aligned} f_{S_1,\ldots,S_{n+1}}(s_1,\ldots,s_{n+1}) &= f_{S_1,\ldots,S_n}(s_1,\ldots,s_n) f_{S_{n+1}|S_1,\ldots,S_n}(s_{n+1}|s_1,\ldots,s_n) && (10)\\ &= \lambda^n \exp(-\lambda s_n) \cdot \lambda \exp(-\lambda(s_{n+1} - s_n)) && (11)\\ &= \lambda^{n+1} \exp(-\lambda s_{n+1}) && (12) \end{aligned}$$

where (11) follows from the fact that given S_n, S_{n+1} is independent of $S_1, \ldots, S_{n-1}$ and furthermore the conditional distribution is that of the inter-arrival time X_{n+1}. This shows that (9) holds.

Now, we would like to obtain $f_{S_n}(s_n)$ from (9). Clearly, we simply integrate out $s_1, \ldots, s_{n-1}$ but we must do so carefully and this is what this note seeks to augment to the textbook. To do so, let us consider something simpler for illustration purposes. Let $n = 3$. Then we are computing

$$\begin{aligned} f_{S_3}(s_3) &= \int_{(s_1,s_2):0<s_1<s_2<s_3} f_{S_1,S_2,S_3}(s_1,s_2,s_3)\, \mathrm{d}s_1\, \mathrm{d}s_2 && (13)\\ &= \int_0^{s_3} \int_0^{s_2} \lambda^3 \exp(-\lambda s_3)\, \mathrm{d}s_1\, \mathrm{d}s_2 && (14)\\ &= \int_0^{s_3} s_2 \lambda^3 \exp(-\lambda s_3)\, \mathrm{d}s_2 && (15)\\ &= \frac{s_3^2}{2} \lambda^3 \exp(-\lambda s_3), \qquad s_3 \in [0,\infty). && (16) \end{aligned}$$

This checks out with (8). For the $n = 4$ case, we have

$$
\begin{aligned}
f_{S_4}(s_4) &= \int_0^{s_4} \int_0^{s_3} \int_0^{s_2} \lambda^4 \exp(-\lambda s_4) \, \mathrm{d}s_1 \, \mathrm{d}s_2 \, \mathrm{d}s_3 && (17) \\
&= \int_0^{s_4} \int_0^{s_3} s_2 \lambda^4 \exp(-\lambda s_4) \, \mathrm{d}s_2 \, \mathrm{d}s_3 && (18) \\
&= \int_0^{s_4} \frac{s_3^2}{2} \lambda^4 \exp(-\lambda s_4) \, \mathrm{d}s_3 && (19) \\
&= \frac{s_4^3}{2 \cdot 3} \lambda^4 \exp(-\lambda s_4) && (20) \\
&= \frac{s_4^3}{3!} \lambda^4 \exp(-\lambda s_4), \qquad s_4 \in [0, \infty). && (21)
\end{aligned}
$$

This checks out with (8). We can easily see a pattern here. The power of s_n always gets incremented and the denominator becomes $(n-1)!$ as we continue integrating. So (8) is true.

More precisely, we have the factor $s_n^{n-1}/(n-1)!$ because the "volume" of the region of the $(s_1, s_2, \ldots, s_{n-1})$ space satisfying $0 < s_1 < s_2 < \ldots < s_{n-1} < s_n$ is exactly $s_n^{n-1}/(n-1)!$, which is exactly what the calculations leading to (16) and (21) are computing.

3 Clarification of Proof 1 of Theorem 2.2.10

The second way to calculate $\Pr(t < S_{n+1} \leq t + \delta)$ was done in a hasty way in the textbook. Let us do it carefully. Define the events

$$
\begin{aligned}
\mathcal{E} &= \{n \text{ arrivals in } (0, t]\} && (22) \\
\mathcal{F} &= \{\text{at least 1 arrival in } (t, t+\delta]\} && (23)
\end{aligned}
$$

Then it is easy to see (draw a picture) that $\Pr(t < S_{n+1} \leq t + \delta) = \Pr(\mathcal{E} \cap \mathcal{F})$. By the independent increments property, events $\mathcal{E}$ and $\mathcal{F}$ are independent so $\Pr(t < S_{n+1} \leq t+\delta) = \Pr(\mathcal{E}) \Pr(\mathcal{F})$. Note, however that $\Pr(\mathcal{E}) = \Pr(N(t) = n) = p_{N(t)}(n)$, the PMF of $N(t)$. On the other hand,

$$
\begin{aligned}
\Pr(\mathcal{F}) &= 1 - \Pr(0 \text{ arrivals in } (t, t+\delta]) && (24) \\
&= 1 - \Pr(0 \text{ arrivals in } (0, \delta]) && (25) \\
&= 1 - \int_\delta^\infty \lambda e^{-\lambda \tau} \, \mathrm{d}\tau && (26) \\
&= \int_0^\delta \lambda e^{-\lambda \tau} \, \mathrm{d}\tau && (27) \\
&= 1 - e^{-\lambda \delta} && (28) \\
&= 1 - [1 - \lambda \delta + o(\delta)] = \lambda \delta + o(\delta), && (29)
\end{aligned}
$$

where importantly, (25) follows from the stationary increments property. In conclusion,

$$
\Pr(t < S_{n+1} \leq t + \delta) = p_{N(t)}(n) \big(\lambda \delta + o(\delta) \big). \tag{30}
$$