Lesson 7. Arrival Counting Processes and the Poisson Arrival Process

1 Overview

- An **arrival counting process** is a stochastic process with one "arrival" system event type
- An "arrival" is broadly defined as any discrete unit that can be counted: for example,
    - customer arrivals
    - service requests
    - accidents in a factory

2 The Case of the Reckless Beehunter

Citizens of Beehunter have complained that a busy intersection has recently become more dangerous, and they are demanding that the city council take action to make the intersection safer. There have been 103 accidents at the intersection since record keeping began. The city council agrees to undertake a study of the intersection to determine if the accident rate has actually increased above the 1 per week average that is (unfortunately) considered normal. It hires a traffic engineer from nearby Vincennes, to perform the study.

The traffic engineer recommends that the number of accidents at the intersection be recorded for a 24-week period. If the number of accidents is significantly larger than expected, then she will declare that the intersection has indeed become more dangerous. During the study period, 36 accidents were observed.

- Our approach:
    - Model the time between accidents as random variables
        - Assume they are independent and time stationary with common cdf F_G (is this reasonable?)
        - $E[G] = 1$ week
    - Model the system as a stochastic process (algorithm that generates possible sample paths)
    - Using this model, determine if the probability that 36 accidents occur in a 24 week period is "small"
- System events:

- State variables:

- System event subroutines:

- Simulation algorithm – the same general algorithm from before:

algorithm Simulation:

1: $n \leftarrow 0$	(initialize system event counter)
$T_0 \leftarrow 0$	(initialize event epoch)
$e_0()$	(execute initial system event)
2: $T_{n+1} \leftarrow \min\{C_1, \dots, C_k\}$	(advance time to next pending system event)
$I \leftarrow \arg\min\{C_1, \dots, C_k\}$	(find index of next system event)
3: $\mathbf{S}_{n+1} \leftarrow \mathbf{S}_n$	(temporarily maintain previous state)
$C_I \leftarrow \infty$	(event I no longer pending)
4: $e_I()$	(execute system event I)
$n \leftarrow n + 1$	(update event counter)
5: go to line 2	

- What does S_n equal for any n?

- Output process: $Y_t \leftarrow S_n$ for all $t \in [T_n, T_{n+1})$, or in words,

- The stochastic process model defined above is called a **renewal arrival-counting process** or a **renewal process** for short
  - arrivals occur one-at-a-time
  - interarrival times are independent and time stationary

3 The Poisson arrival process

- Based on historical data and some statistical testing, the traffic engineer from Vincennes has determined that the time between accidents are in fact exponentially distributed with a mean of 1 week

⇒ Let's now assume that $G \sim \text{Exponential}(\lambda)$:

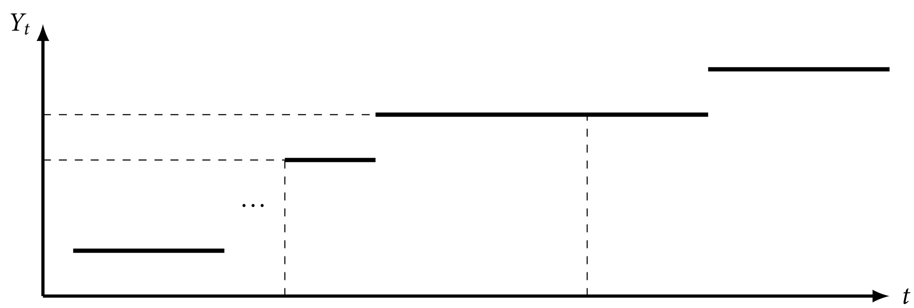
$$F_G(a) = 1 - e^{-\lambda a} \quad \text{for } a \geq 0 \qquad E[G] = \frac{1}{\lambda} \qquad \text{Var}(G) = \frac{1}{\lambda^2}$$

- This type of renewal process is known as a **Poisson arrival process**
- Let G_i be the interarrival time between arrivals $i-1$ and i
- We can directly write the event epoch T_n as a function of the interarrival times $G_1, \ldots, G_n$:

- Since G_i are exponential random variables with parameter λ, T_n is an **Erlang random variable with parameter λ and n phases**:

$$F_{T_n}(a) = 1 - \sum_{j=0}^{n-1} \frac{e^{-\lambda a}(\lambda a)^j}{j!} \quad \text{for } a \geq 0 \qquad E[T_n] = \frac{n}{\lambda} \qquad \text{Var}(T_n) = \frac{n}{\lambda^2}$$

- The output process $Y_1, Y_2, \ldots$ and the event epoch process $T_1, T_2, \ldots$ are fundamentally related:



- Therefore, we can get an explicit expression for the cdf of Y_t:

- And we can also get an expression for the pmf of Y_t:

- The pmf and cdf may look familiar: Y_t is a **Poisson random variable** with parameter λt

$$\Rightarrow \quad E[Y_t] = \lambda t \qquad \text{Var}(Y_t) = \lambda t$$

Example 1. In the Beehunter case, the inter-accident times were exponentially distributed with parameter $\lambda = 1$. What is the probability that the total number of accidents at week 24 is greater than 36? What is the expected number of accidents? (Your calculator can compute summations.)

4 Properties of the Poisson process

- Let $\Delta t > 0$ be a time increment
- The **forward-recurrence time** R_t is the time between t and the next arrival
- The **independent-increments property**: the number of arrivals in nonoverlapping time intervals are independent random variables:
  - As a consequence:
- The **stationary-increments property**: the number of arrivals in a time increment of length Δt only depends on the length of the increment, not when it starts:
  - As a consequence:

  $\Rightarrow$ λ can be interpreted as the **arrival rate** of the Poisson process
- The **memoryless property**: the forward-recurrence time R_t has the same distribution as the interarrival time:
- These properties make computing probability statements about Poisson processes pretty easy

Example 2. Recall that in the Beehunter case, a total of 103 accidents have occurred at the intersection up to the time the traffic engineer starts observing, time a. What is the probability that more than 36 accidents are observed in the following 24 weeks? What is the expected number of accidents?

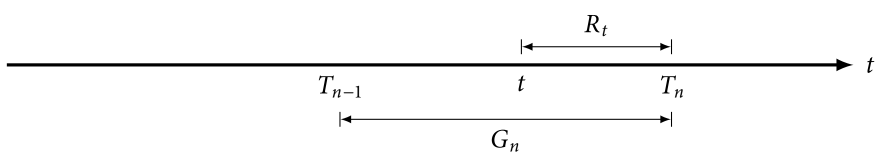
Example 3. What is the probability there are 4 accidents at week 5, given that there are 2 accidents at week 4? What is the expected number of accidents at week 5?

Example 4. What is the probability that the 10th arrival occurs before the 7th week? What is the expected time of the 10th arrival?

- **Any arrival-counting process in which arrivals occur one-at-a-time and has independent and stationary increments must be a Poisson process**
    - If you can justify your system having independent and stationary increments, then you can assume that interarrival times are exponentially distributed
    - This is a very deep and powerful result

5 Why does the memoryless property hold?

- The memoryless property allows us to ignore when we start observing the Poisson process, since forward-recurrence times and interarrival times are distributed in the same way
- "Memoryless" $\longleftrightarrow$ how much time has passed doesn't matter
- Why is this true for Poisson processes?
- Let's consider G_n, the interarrival time between the $(n-1)$th and nth arrival (between T_{n-1} and T_n)
    - Recall that $G_n \sim \text{Exponential}(\lambda)$
- Pick some t between T_{n-1} and T_n
- We want to show that the forward-recurrence time $R_t \sim \text{Exponential}(\lambda)$
    - Equivalently, we show $F_{R_t}(a) = \Pr\{R_t \le a\} = 1 - e^{-\lambda a}$



- Therefore:

$$
\begin{aligned}
\Pr\{R_t > a\} &= \Pr\{G_n > t - T_{n-1} + a \mid G_n > t - T_{n-1}\} \\
&= \frac{\Pr\{G_n > t - T_{n-1} + a \ \text{ and } \ G_n > t - T_{n-1}\}}{\Pr\{G_n > t - T_{n-1}\}} \\
&= \frac{\Pr\{G_n > t - T_{n-1} + a\}}{\Pr\{G_n > t - T_{n-1}\}} \\
&= \frac{e^{-\lambda(t - T_{n-1} + a)}}{e^{-\lambda(t - T_{n-1})}} = e^{-\lambda a}
\end{aligned}
$$

$$\Rightarrow \quad \Pr\{R_t \le a\} = 1 - e^{-\lambda a}$$

- Note: This "proof" is rough and sketchy – we actually need to condition on T_{n-1} and Y_t
    - Repeated use of the law of total probability
- The independent-increments and stationary-increments properties follow from the memoryless property and the fundamental relationship between Y_t and T_n (see Nelson pp. 110-111)