Lesson 9. Introduction to Poisson Processes

1 Overview

- In this lesson: a new kind of stochastic process that models arrivals to a system
- An "arrival" is broadly defined as any discrete unit that can be counted, for example:
  - customer arrivals
  - service requests
  - accidents at an intersection

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. 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. Based on historical data, she has determined that the time between accidents is exponentially distributed with a mean of 1 week, implying that the average number of accidents during a 24-week period is 24. During the study period, 36 accidents were observed. Is 36 accidents "significantly" more than what we expect?

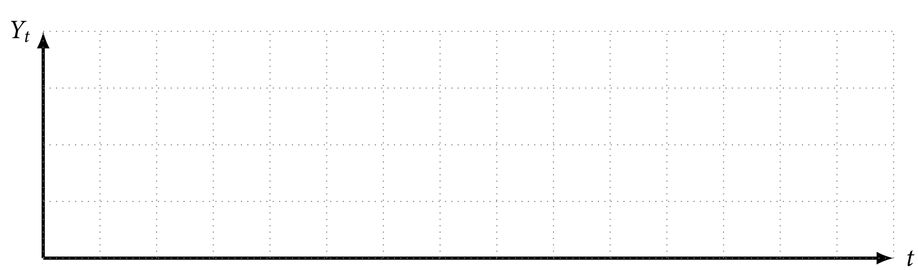
3 Arrival counting processes

- Suppose that "arrivals" come to a system one at a time
- Let

$$\begin{aligned} G_n &= \text{the } n\text{th } \textbf{interarrival time}\text{, or the time between the } (n-1)\text{th arrival and the } n\text{th arrival} \\ T_n &= \text{the } n\text{th } \textbf{arrival time}\text{, or the time of the } n\text{th arrival} \\ Y_t &= \text{the number of arrivals by time } t \end{aligned}$$

- Furthermore, we assume the system is empty at time 0:
- $\{Y_t : t \geq 0\}$ is an **arrival counting process**
  - Continuous-time, discrete-state stochastic process

Example 1. Suppose the outputs of G_1, G_2, and G_3 are 3, 2, and 4, respectively. Graph Y_t.



- We can write T_n in terms of the interarrival times $G_1, \ldots, G_n$:

- The number of arrivals Y_t and the arrival times T_n are fundamentally related:

4 The Poisson process

- Now, let's further assume that $G_1, G_2, \cdots \sim$ iid Exponential(λ)
  - Note that the interarrival times are **time stationary**: the distribution stays the same over time

$\Rightarrow$ $\{Y_t : t \geq 0\}$ is a **Poisson arrival process with arrival rate** λ

  - Why "Poisson" and "arrival rate"? We'll see soon.

- Recall that

$$G_1, G_2, \ldots G_n \sim \text{iid Exponential}(\lambda) \Longrightarrow T_n \sim \text{Erlang}(n, \lambda) \qquad F_{T_n}(a) = 1 - \sum_{k=0}^{n-1} \frac{e^{-\lambda a}(\lambda a)^k}{k!} \quad \text{for } a \geq 0$$

- Therefore, the cdf of Y_t is:

- And so the pmf of Y_t is:

- The pmf and cdf look familiar...

- Therefore, the average number of arrivals by time t is

- So, the average **arrival rate** in this arrival process is

Example 2. 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?

5 Properties of the Poisson process

- Let $\Delta t > 0$ be a time increment

- 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:

- The **forward-recurrence time** R_t is the time between t and the next arrival
- 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 3. 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 4. 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 5. What is the probability that the 10th arrival occurs before the 7th week? What is the expected time of the 10th arrival?

6 When is the Poisson process a good model?

- **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 arrivals having independent and stationary increments, then you can assume that the interarrival times are exponentially distributed
  - This is a very powerful result
- Independent increments ⇔ number of arrivals in nonoverlapping intervals of time are independent
  - Reasonable when the arrival-counting process is formed by a large number of customers making individual, independent decisions about when to arrive
- Stationary increments ⇔ expected number of arrivals = constant rate × length of time interval
  - Reasonable when arrival rate is approximately constant over time

Example 6. Discuss whether or not it is reasonable to approximate the following arrival processes as Poisson processes:

a. The arrival of cars at a toll booth during evening rush hour.
b. The arrival of students at a college football game.

7 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 $F_{R_t}(a) = \Pr\{R_t \leq a\} = 1 - e^{-\lambda a}$

- Therefore:

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

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

- 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 SMAS pp. 110-111)

8 Exercises

Problem 1 (SMAS Exercise 5.1)**.** For a Poisson process $\{Y_t : t \geq 0\}$ with arrival rate $\lambda = 2$ per hour, compute the following:

a. $\Pr\{Y_2 = 5\}$
b. $\Pr\{Y_4 - Y_3 = 1\}$
c. $\Pr\{Y_6 - Y_3 = 4 \mid Y_3 = 2\}$
d. $\Pr\{Y_5 = 4 \mid Y_4 = 2\}$

Problem 2 (SMAS Exercise 5.3)**.** Patients arrive at a hospital emergency room at a rate of 2 per hour. A doctor works a 12-hour shift from 6 a.m. until 6 p.m. Answer the following questions by approximating the arrival-counting process as a Poisson process.

a. If the doctor has seen 6 patients by 8 a.m., what is the probability that the doctor will see a total of 9 patients by 10 a.m.?
b. What is the expected time between the arrival of successive patients? What is the probability that the time between the arrival of successive patients will be more than 1 hour?
c. What is the expected time after coming on duty until the doctor sees her first patient? What is the probability that she will see her first patient in 15 minutes or less after coming on duty?
d. What is the probability that the doctor will see her thirteenth patient before 1 p.m.?

Problem 3 (SMAS Exercise 5.6)**.** An automated optical scanner looks for defects on a continuous sheet of metal. If the metal is being produced according to specifications, then defects should occur at a rate of 1 defect per 50 square meters of metal. Assume that the occurrence of defects can be modeled as a Poison arrival process. Notice that "time" corresponds to square meters of metal in this situation. What is the probability of finding 7 or more defects in any 200 m^2 of metal?

Problem 4. The Markov Company has a manufacturing cell that processes jobs during a 12-hour shift starting at 6 a.m. and ending at 6 p.m. Jobs leave the cell according to a Poisson process with rate $\lambda = 8$ per hour.

a. If the cell has processed exactly 10 jobs by 8 a.m., what is the probability that the cell will have processed more than 30 jobs by 10 a.m.?
b. What is the probability that the cell will have processed its 50th job before 12 p.m.?
c. If the cell has processed exactly 40 jobs by 12 p.m., what is the probability that the cell will have processed its 100th job by the end of the shift?
d. When are the first 4 jobs expected to be completed? (Assume all jobs are available starting at 6 a.m.)

Problem 5 (SMAS Exercise 5.14)**.** Suppose that the occurrence of typographical errors in the first draft of a book is well approximated as a Poison process with a rate of 1 error per 1000 words. Each time the author proofreads the book, the error rate is reduced by 50%. For example, after the first proofreading, the error rate is 0.5 errors per 1000 words, or 1 error per 2000 words. Notice that we are saying that the "error rate" goes down by 50%, not that the actual number of errors is reduced by that amount since the author can (and did) introduce new errors while correcting other errors. The author wants to know how many times he needs to proofread the book so that the probability of there being no errors is at least 0.98. Assume that the book contains 200,000 words. Compute this number.

Problem 6 (SMAS Exercise 5.8)**.** Several arrival processes are described below. State whether or not it is reasonable to approximate them as Poisson processes, and why or why not.

a. The arrival of customers to the purchase doughnuts from a campus doughnut shop.
b. The arrival of students at a university football game or concert.
c. The arrival of patients at a doctor's office.
d. The discovery of bugs in a new software product.
e. The arrival of emergency calls to a fire station.