Lecture notes 13

Stochastic Processes I: Bernoulli and Poisson Processes

Lecture Outline

- Stochastic Processes
- Bernoulli Process
- Binomial process
- Interarrival times
- Poisson process
- Random telegraph wave
- Arrivals

- Interarrival times
- Properties

Reading: Bertsekas & Tsitsiklis 5.1, 5.2

Stochastic Processes

Both mathematically and intuitively, a *stochastic process* or *random process* is simply an infinite collection of random variables, one for each value of time (or, in some examples, space)

For example, an iid sequence of Bernoulli random variables $\{X_n; n = 0, 1, \ldots, \}$ that goes on forever is a *Bernoulli stochastic process*, the mathematical model often used to model real phenomena like flipping a biased coin or a sequence of binary data.

Usually used for binary processes, but name is used more generally for any iid sequence of random variables, e.g., rolls of a die.

When the *index* of the random variables takes on discrete values (e.g., X_n with n contained in the integers or the nonnegative integers), then call it a *discrete time stochastic process*

Also have *continuous time stochastic processes* such as $\{X(t);\ t \geq 0\}$ or $\{X_t;\ t \in \Re\}$, where all of the $X(t)$ are random variables defined on a common experiment. We will see an example shortly.

Examples of stochastic processes

- scores of consecutive NFL games
- daily prices of a stock
- thermal noise in a resistor
- numbers generated by successive spinning of a roulette wheel
- winnings or losses of a roulette player

Important Issues

- Dependence or independence of variables in a stochastic process, does knowledge of the past help predict the future?
  - For a Bernoulli process knowing the past does not help.
  - For a Markov chain knowing about the past can help a lot.
- Long term averages, e.g.,
  - mean power through a line
  - proportion of time a queue is empty
  - average number of machine failures per day
- Extremes
  - What is the probability that all circuits at an exchange will be busy for a full hour?
  - What is the probability that a maximum allowable power threshold will be exceeded?

Bernoulli Process

Assume that $\{X_n;\ n = 1, \ldots\}$ is an infinite sequence of the random variables with the property that if we "sample" any finite number of the random variables, we get a finite collection of iid Bernoulli random variables with parameter p.

$$E(X_n) = p \quad , \quad \sigma^2_{X_n} = p(1-p) \text{ all } n$$

We say that $\{X_n;\ n = 1, \ldots\}$ is a *Bernoulli random process* or an infinite sequence of Bernoulli trials.

- Infinite sequence of coin flips
- Sequence of binary symbols on a digital computer link with no known ending point
- Pick a real number at random on the unit interval and print its binary expansion

Properties of Bernoulli Process

- Memoryless property:
  - Any events defined on nonoverlapping sets of indexes (times, trials) are independent.
- Fresh-start property:
  - Starting with time n, the process behaves identically (in a probabilistic sense) as starting with time 1, regardless of what happened in the past.

IID Processes

Generalization of a Bernoulli process: A process $\{X_n\}$ is *independent and identically distributed* (IID, iid, i.i.d.) if any finite collection of random variables drawn from the sequence is mutually independent and all cdfs F_{X_n} are the same, say F_X.

Any pmf p_X or pdf f_X can be used to define an IID process.

For example, if f_X is a Gaussian pdf, then the resulting process is a Gaussian IID process, also called *discrete time Gaussian white noise*.

Number of Successes

Recall that if fix n and define $S_n = \sum_{i=1}^{n} X_i$ = number of successes in n Bernoulli trials, then S_n has a binomial (n, p) pmf

$$\begin{aligned} p_{S_n}(k) &= \binom{n}{k} p^k (1-p)^{n-k},\ k = 0, 1, \ldots, n \\ E[S_n] &= np \\ \sigma^2_{S_n} &= np(1-p) \\ M_{S_n}(s) &= ((1-p) + pe^s)^n \end{aligned}$$

$\{S_n;\ n = 1, 2, \ldots\}$ is itself a random process, called the *binomial counting process* (it counts ones/successes in the underlying Bernoulli process)
S_n *is not iid!!*

For very large n, very small p, and moderate $\lambda = np$, the binomial pmf can be approximated by the Poisson pmf (See B& T)

$$\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}$$

Example

- A person plays the lottery every day.
- Each ticket has probability $p = 0.001$ of winning
- How many tickets does the person need to buy in order to have at least a 10% chance of winning?

If n tickets are bought, need

$$P(\text{ win }) = 1 - p_S(0) = 1 - (0.999)^n \geq 0.1$$

or

$$n \geq \frac{\ln 0.9}{\ln 0.999} = 105.3605$$

Using the Poisson approximation, with $\lambda = .001n$

$$P(\text{ win }) = 1 - p_S(0) \approx 1 - e^{-0.001n} \geq 0.1$$

also get $n \geq 105.3605$

Interarrival Times

Consider a Bernoulli process $\{X_n;\ n = 1, 2, \ldots\}$

Let Y_1 denote the number of trials (starting at $n = 1$) needed for the first success, i.e., $Y_1 = k$ if $X_1 = X_2 = \cdots X_{k-1} = 0$ and $X_k = 1$.

Have seen that Y_1 has a geometric pmf with parameter p, i.e.,

$$\begin{aligned} p_{Y_1}(k) &= p(1-p)^{k-1};\ k = 1, 2, \ldots \\ E[Y_1] &= \frac{1}{p} \\ \sigma^2_{Y_1} &= \frac{1-p}{p^2} \\ M_{Y_1}(s) &= \frac{pe^s}{1-(1-p)e^s} \end{aligned}$$

From the memoryless property of the Bernoulli process, Y_1 could be defined as the number of trials to the next success starting at any time. E.g., if you are told that $X_m = 1$, then the length of time to the *next* success, i.e., the smallest value of k for which $X_{m+1} = X_{m+2} = \cdots X_{m+k-1} = 0$ and $X_{m+k} = 1$ has the same pmf as Y_1.

Y_1 called an *interarrival time*, the time between successive successes (or 1's or "arrivals")

Example Using the previous lottery example with $p = .001$, if the person buys 1 ticket per day, what is the expected length of strings of losing days?

$$\begin{aligned} E[\text{ length }] &= E[Y_1] \\ &= \frac{1}{p} \\ &= 1000 \end{aligned}$$

Interarrival Times of Order > 1

Now let Y_k be the number of trials to the kth success

Y_k will be the sum of k iid geometric RVs (because the underlying Bernoulli process is memoryless), so

$$\begin{aligned} E[Y_k] &= \frac{k}{p} \\ \sigma^2_{Y_k} &= \frac{k(1-p)}{p^2} \\ M_{Y_k}(s) &= \left(\frac{pe^s}{1-(1-p)e^s}\right)^k \end{aligned}$$

It can be shown that the inverse transform of $M_{Y_k}(s)$ s

$$p_{Y_k}(t) = \binom{t-1}{k-1} p^k (1-p)^{t-k};\ t = k, k+1, k+2, \ldots; k = 1, 2, \ldots$$

Called the *Pascal* pmf of order k

Note that $\{Y_n; n = 1, 2, 3, \ldots\}$ is yet another random process.

The Poisson Process

- Poisson process involved with events or arrivals, like 1's in a Bernoulli process. With a Poisson process things can happen continuously in time.
- Examples include customer arrivals at a bank, email receptions, packet arrivals at a switch, busses arriving at a bus stop, phone calls arriving at an exchange
- A Bernoulli process can model arrivals in discrete time. Will see that a Bernoulli process resembles a Poisson process with finely discretized time.

Suppose we have a random process $\{X(t)\ 0 \leq t < \infty\}$ which like the Bernoulli process can only take on values of 0 or 1. Points in time where $X(t)$ changes values will be called *arrivals*.

Define $p(k,t)$ as the probability there are k arrivals in an interval of duration t, where $t \geq 0$ and $k = 0, 1, \ldots$.

Note that t is a parameter and $p(k,t)$ is a pmf.

Can make this look more familiar by defining a RV for any positive real t as $N_t =$ number of arrivals in $[0,t)$.

Then $p(k,t) = p_{N_t}(k)$, the pmf for the RV N_t.

Thus N_t counts the arrivals or changes or jumps in X_t.

$\{N_t;\ t \geq 0\}$ is called a *counting process*

Note: The sum of iid Bernoulli random variables is also a counting process, i.e., a process that increases by increments of 1 according to the occurrence of certain events.

Make the following assumptions about the process:

- $p(k,t)$ is the same for any interval of length t (like the Bernoulli process).
- the number of arrivals in disjoint time itervals are mutually independent random variables (like the Bernoulli process)
- If Δt is a differentially small interval then

$$p(k,\Delta t) \approx \begin{cases} 1-\lambda\Delta t & \text{if } k=0 \\ \lambda\Delta t & \text{if } k=1 \\ 0 & \text{if } k>1 \end{cases}$$

What is $p(k,t)$?

Use conditional probability to evaluate the number of arrivals in the interval $[0,t+\Delta t)$ in terms of the number of arrivals in the left part of the interval $[0,t)$: $[0,t)$ $[t,t+\Delta t)$

$[0,t)$ $[t,t+\Delta t)$

$$\begin{aligned}
&p(k,t+\Delta t)\\
&\quad= P(k \text{ arrivals in } [0,t+\Delta t))\\
&\quad= \sum_{l=0}^{k} P(l \text{ arrivals in } [t,t+\Delta t) \mid k-l \text{ arrivals in } [0,t))\\
&\qquad\quad \times P(k-l \text{ arrivals in } [0,t))\\
&\quad= \sum_{l=0}^{k} P(l \text{ arrivals in } [t,t+\Delta t)) \times P(k-l \text{ arrivals in } [0,t))\\
&\quad\approx P(0 \text{ arrivals in } [t,t+\Delta t)) \times P(k \text{ arrivals in } [0,t))\\
&\qquad\quad + P(1 \text{ arrival in } [t,t+\Delta t)) \times P(k-1 \text{ arrivals in } [0,t))\\
&\quad= p(0,\Delta t)p(k,t) + p(1,\Delta t)p(k-1,t)\\
&\quad\approx (1-\lambda\Delta t)p(k,t) + \lambda\Delta t p(k-1,t)
\end{aligned}$$

A little algebra yields

$$\frac{p(k,t+\Delta t)-p(k,t)}{\Delta t} = p(k-1,t)\lambda - p(k,t)\lambda \ .$$

In the limit as $\Delta t \to 0$ this becomes the differential equation

$$\frac{d}{dt}p(k,t) + \lambda p(k,t) = \lambda p(k-1,t) \ , \ t > 0 \ .$$

As an initial condition for this differential equation assume

$$p(k,0) = \begin{cases} 0 & , \quad k \neq 0 \\ 1 & , \quad k = 0 \ , \end{cases}$$

The solution to the differential equation with the given initial condition is

$$p(k,t) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}; \ k = 0,1,2,\ldots; \ t \geq 0$$

This is easily verified by direct substitution:

$$\begin{aligned} \frac{d}{dt}p(k,t) &= \frac{1}{k!}\left(k\lambda^k t^{k-1}e^{-\lambda t} + (\lambda t)^k(-\lambda)e^{-\lambda t}\right) \\ &= \lambda\frac{(\lambda t)^{k-1}e^{-\lambda t}}{(k-1)!} - \lambda p(k,t) \\ &= \lambda p(k-1,t) - \lambda p(k,t) \\ p(k,0) &= \begin{cases} 0 & k = 0 \\ 1 & \text{otherwise} \end{cases} \end{aligned}$$

A Poisson pmf!

Note that this pmf indeed satisfies the underlying assumptions, i.e.,

$$p(k,\Delta t) \approx \begin{cases} 1 - \lambda\Delta t & \text{if } k = 0 \\ \lambda\Delta t & \text{if } k = 1 \\ 0 & \text{if } k > 1 \end{cases}$$

This provides another means of defining a Poisson process: A Poisson process is one for which

- the number of arrivals in an interval depends only on the length of the interval
- arrivals in nonoverlapping time intervals are independent, and
- $$p(k,t) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}; \ k = 0,1,2,\ldots; \ t \geq 0$$

- The process $\{X(t);\ t \geq 0\}$ is called the *random telegraph wave* and it was a much studied process in the early days of telecommunications (especially by Norbert Weiner)
- The process $\{N_t;\ t \geq 0\}$ is called the *Poisson counting process*

Have shown

$$p_{N_t}(k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}; \ k = 0,1,2,\ldots; \ t \geq 0$$

Hence know moments and transform:

$$E[N_t] = \lambda t \qquad \sigma^2_{N_t} = \lambda t \qquad M_{N_t}(s) = e^{\lambda t(e^s - 1)}$$

λ is called the *arrival rate*

Example You get email according to a Poisson process at a rate of $\lambda = 0.2$ messages per hour. You check your email every hour. What is the probability of finding 0 and 1 new messages?

$$\begin{aligned} p(0,1) &= e^{-0.2} \\ &= 0.819 \\ p(1,1) &= 0.2e^{-0.2} \\ &= 0.164 \end{aligned}$$

Suppose that you have not checked your email for a full day. What is the probability of finding no new messages?

$$p(0,24) = e^{-0.2\times 24} \approx 0.008294$$

Interarrival Times

So far fixed a time t and found the pmf for the number of arrivals in an interval of length t

Now reverse the problem: Suppose fix the number of arrivals as k and ask for the pdf of the time τ_k of the kth arrival.

E.g., consider consider $k = 1$ and find the pdf of τ_1.

As usual, first find the cdf:

$$\begin{aligned} F_{\tau_1}(t) &= P(\text{ first arrival occurs before } t\ ) \\ &= P(N_t \geq 1) \\ &= 1 - p_{N_t}(0) \\ &= 1 - e^{-\lambda t};\ t \geq 0 \end{aligned}$$

Taking derivatives to get the pdf yields

$$\begin{aligned} f_{\tau_1}(t) &= \frac{d}{dt}F_{\tau_1}(t) \\ &= \lambda e^{-\lambda t};\ t \geq 0 \end{aligned}$$

An exponential pdf!

For the kth order interarrival times τ_k, can argue that it is the sum of k iid exponential RVs and invert the transform.

E.g., suppose that $T_1 = \tau_1 =$ time of the first arrival and $T_i = \tau_i - \tau_{i-1}$ $=$ first order interarrival time between $(i-1)$th arrival and ith arrival for $i > 1$. Then the T_i are exponentially distributed and independent (since they depend on disjoint intervals) and

$$\tau_k = \sum_{i=1}^{k} T_i \quad \Rightarrow \quad M_{\tau_k}(s) = (\frac{\lambda}{\lambda - s})^k$$

More direct:

$$\begin{aligned} F_{\tau_k}(t) &= P(\ k\text{th arrival occurs before } t\ ) \\ &= P(N_t \geq k) \\ &= \sum_{l=k}^{\infty} \frac{(\lambda t)^l e^{-\lambda t}}{l!} \\ &= 1 - \sum_{l=0}^{k-1} \frac{(\lambda t)^l e^{-\lambda t}}{l!} \\ &= 1 - e^{-\lambda t} - \sum_{l=1}^{k-1} \frac{(\lambda t)^l e^{-\lambda t}}{l!} \end{aligned}$$

Taking derivatives to get the pdf yields

$$
\begin{aligned}
f_{\tau_k}(t) &= \frac{d}{dt}F_{\tau_k}(t) = \frac{d}{dt}\left(1 - e^{-\lambda t} - \sum_{l=1}^{k-1}\frac{(\lambda t)^l e^{-\lambda t}}{l!}\right) \\
&= \lambda e^{-\lambda t} - \sum_{l=1}^{k-1}\frac{d}{dt}\frac{(\lambda t)^l e^{-\lambda t}}{l!} \\
&= \lambda e^{-\lambda t} - \sum_{l=1}^{k-1}\frac{\lambda^l l t^{l-1} e^{-\lambda t}}{l!} + \sum_{l=1}^{k-1}\lambda\frac{(\lambda t)^l e^{-\lambda t}}{l!} \\
&= \lambda e^{-\lambda t} - \lambda\sum_{l=1}^{k-1}\frac{(\lambda t)^{l-1} e^{-\lambda t}}{(l-1)!} + \lambda\sum_{l=1}^{k-1}\frac{(\lambda t)^l e^{-\lambda t}}{l!} \\
&= \lambda e^{-\lambda t} - \lambda\sum_{l=0}^{k-2}\frac{(\lambda t)^l e^{-\lambda t}}{(l)!} + \lambda\sum_{l=1}^{k-1}\frac{(\lambda t)^l e^{-\lambda t}}{l!} \\
&= \frac{\lambda^k t^{k-1} e^{-\lambda t}}{(k-1)!};\ t \geq 0;\ k = 1, \ldots
\end{aligned}
$$

The *Erlang* family of pdf's.

Using the fact that τ_k is the sum of k iid exponential random variables,

$$
\begin{aligned}
E[\tau_k] &= \frac{k}{\lambda} \\
\sigma^2_{\tau_k} &= \frac{k}{\lambda^2} \\
M_{\tau_k}(s) &= M_{\tau_1}(s)^k \\
&= (\frac{\lambda}{\lambda - s})^k
\end{aligned}
$$

RV's Associated with Poisson Processes

- Poisson pmf with parameter (arrival rate) λ:

$$
\begin{aligned}
p_{N_t}(k) &= \frac{(\lambda t)^k e^{-\lambda t}}{k!};\ k = 0, 1, 2, \ldots;\ t \geq 0 \\
E[N_t] &= \lambda t \\
\sigma^2_{N_t} &= \lambda t \\
M_{N_t}(s) &= e^{\lambda t(e^s - 1)}
\end{aligned}
$$

- Exponential pdf with parameter λ: $T = \tau_1 = T_1 =$ first order interarrival times:

$$
f_T(t) = \lambda e^{-\lambda t};\ t \geq 0
$$

$$
\begin{aligned}
E[T] &= \frac{1}{\lambda} \\
\sigma^2_T &= \frac{1}{\lambda^2} \\
M_T(s) &= (\frac{\lambda}{\lambda - s})
\end{aligned}
$$

- Erlang pdf of order k and parameter λ: $\tau_k = k$th order interarrival times

$$
\begin{aligned}
f_{\tau_k}(t) &= \frac{\lambda^k t^{k-1} e^{-\lambda t}}{(k-1)!};\ t \geq 0;\ k = 0, 1, \ldots \\
E[\tau_k] &= \frac{k}{\lambda} \\
\sigma^2_{\tau_k} &= \frac{k}{\lambda^2} \\
M_{\tau_k}(s) &= M_{\tau_1}(s)^k \\
&= (\frac{\lambda}{\lambda - s})^k
\end{aligned}
$$

	Poisson	Binomial
arrival times	continuous	discrete
pmf of # of arrivals	Poisson	Binomial
Interarrival times	exponential	geometric
arrival rate	λ/unit time	p/unit time (trial)
kth interarrival times	Erlang	Pascal

Connection: Suppose have a Poisson process and discretize (quantize) time into tiny intervals $[n\Delta t, (n+1)\Delta t)$ $n = 0, 1, \ldots$. Then the probability of an arrival in the nth interval is $p \approx \lambda\Delta t$ and the probability of no arrival is $1 - p$ and the probability of more than one arrival is approximately 0.

Thus time-quantized Poisson behaves approximately like a Bernoulli process.

(Binomial is approximately Poisson when parameters chosen suitably, see book.)

Example

- You call the IRS hotline and you are told you are the 56th person in line. Callers depart according to a Poisson process with rate $\lambda = 2$ per minute.
  - How long will you have to wait on the average?
  - What is the probability that you will have to wait for more than an hour?

Memoryless property implies service time to kth customer is distributed as τ_k and hence has mean $E[\tau_k] = 56/\lambda = 28$.

The probability you have to wait for more than an hour is given by the formula

$$P(\tau_k \geq 60) = \int_{60}^{\infty} f_{\tau_k}(t)\,dt = \int_{60}^{\infty} \frac{\lambda^k t^{k-1} e^{-\lambda t}}{(k-1)!}\,dt$$

which is a particularly nasty integral. Need to do it numerically or approximately using the central limit theorem (to be described later)

Adding Poisson RV's and Processes

Suppose that N and L are independent Poisson random variables with rates λ_1 and λ_2, respectively.

Define $K = N + L$. Derive the pmf of K.

As always, this is easiest with transforms:

$$\begin{aligned} M_K(s) &= E[e^K] \\ &= E[e^{N+L}] = E[e^N]E[e^L] \\ &= M_N(s)M_L(s) = e^{\lambda_1 t(e^s-1)}e^{\lambda_2 t(e^s-1)} \\ &= e^{(\lambda_1+\lambda_2)t(e^s-1)} \end{aligned}$$

i.e., K is also Poisson and its rate is $\lambda_1 + \lambda_2$:

$$p_K(k) = \frac{((\lambda_1+\lambda_2)t)^k e^{-(\lambda_1+\lambda_2)t}}{k!};\ k = 0, 1, \ldots$$

Now suppose have two mutually independent random processes with rates λ_1 and λ_2. Create a new process, where an arrival occurs when an arrival from process 1 or process 2 occurs.

Let N_t denote the counts of the first process and L_t denote the counts of the second. Have seen that $K_t = N_t + L_t$ will then also be Poisson with a rate equal to the sum of the rates. Furthermore, arrivals in nonoverlapping intervals of the new process will be independent.

Thus the sum or "merged" Poisson processes will yield a new Poisson process!