MA 225 Second Half; Stochastic Processes (Ross)

October–November 2021

Contents

1 WARNING!

- This document **DOES NOT** contain the full content of the lectures, and **SHOULD NOT** be regarded as a substitute for the blackboard lecture.
- This document is under preparation. Much of material of the lectures is yet to be included. The included material is also likely to contain typos and other similar errors.
- I will try to share an updated version if ther eare significant additions or alterations. So please share any typos that you might find. Thanks!

2 Week one

`Tuesday 28 September to Thursday 30 September 2021`

3 Week two: roll a die till you get ten even outcomes

`Monday 4 October to Thursday 7 October 2021`

1. `Tutorial on Monday 4 October`. Subrata solved the mid semester questions.

2. Showing $E[X_1] = E[X_2]$ is similar in spirit to the following analogous question for coin tosses. The analogous setting has the virtue of stripping away unnecessary complexity.

   `Question`. Toss a coin till you get 10 heads. Let N_{10} denote the number of tosses it takes to get 10 heads. Let $T_{N_{10}}$ denote the number of tails in N_{10} coin tosses. Show that $E[T_{N_{10}}] = 10$.

   `Answer`. `PENDING`: First state strong law of large numbers. By the strong law of large numbers, almost surely $\frac{H_n}{n} \to \frac{1}{2}$ and $\frac{T_n}{n} \to \frac{1}{2}$, and so $\frac{T_n}{H_n} \to 1$ almost surely as $n \to \infty$. Since the limit must remain the same along any subsequence (whether deterministic or random), we get

   $$\frac{T_{N_{10k}}}{H_{N_{10k}}} \to 1 \text{ almost surely as } k \to \infty.$$

   Note that $T_{N_{10k}}$ can be regarded as a sum of k i.i.d. random variables, each distributed like $T_{N_{10}}$. Further, since $H_{N_{10k}} = 10k$, we get

   $$\frac{\text{Sum of } k \text{ i.i.d. random variables, each distributed like } T_{N_{10}}}{k} \to 10.$$

   By the strong law of large numbers the limit must be $E[T_{N_{10}}]$. Combining, we get

   $$E[T_{N_{10}}] = 10$$

3. `Question`. How is the random variable X_1 distributed?

   `Answer`. Here, we are only interested in those rolls of the die that yield either a "1″ or an even outcome. We can do this by (iteratively) replacing all occurrences of "3″ and "5″ by a repeat roll of the die, in which case we effectively get only one of four outcomes $\{1, 2, 4, 6\}$, each with an equal probability of $\frac{1}{4}$. In this interpretation, if $X_1 = k$ then the 10^{th} occurrences of an even number must have occurred on the $(k+10)^{\text{th}}$ roll. So the first $k+9$ rolls must have 9 occurrences of an even number. We get

   $$P\{X_1 = k\} = \binom{k+9}{9}(1/4)^k(3/4)^{10}.$$

4. `Question`. How is the random variable X_2 distributed?

   `Answer`. There are 10 occurrences of an even number, and each occurrence is equally likely to be any of the three even numbers available on a die.

   $$P\{X_2 = k\} = \binom{10}{k}(1/3)^k(2/3)^{10-k}.$$

5.
   - `Question.` Are Y_i i.i.d.?

     `Answer.`

- `Question.` Are X_1, X_3, X_5 i.i.d.?
  `Answer.`
- `Question.` Are X_2, X_4, X_6 i.i.d.?
  `Answer.`
- `Question.` Are X_4, X_6 i.i.d.?
  `Answer.`

4 Week three: SLLN, linear graph, first failure, infection time

`Monday 11 October to Thursday 14 October 2021`

1. `Tutorial on Monday 11 October`. Mijanur solved the next two questions. Subsequent material done in lectures.

2. `Problem 1.37, Stochastic Processes, Ross`. Let (X_i) be a sequence of i.i.d. continuous random variables. X_n is said to be a peak if it is greater than both its neighbours. Show that almost surely the fraction of variables that are a peak tends to $\frac{1}{3}$.

   `Answer.` Work with indicator random variables and apply strong law of large numbers thrice to ensure independence of terms.

3. `Problem 1.39, Stochastic Processes, Ross`. A graph consists of vertices v_0 to v_n connected in a line. Show that expected number of steps to go from v_0 to v_n is n^2.

   `Answer.` Let T_k be the number of steps between first visit to v_{k-1} and first visit to v_k. Let

   $$T = \sum_1^n T_k.$$

   To show $E[T] = n^2$, it is sufficient to show $E[T_k] = k^2 - (k-1)^2 = 2k - 1$. We prove this by induction. First note that $T_1 = 1$. We assume $T_l = 2l - 1$ for all $l < k$ and extend it to $T_k = 2k - 1$. We start by considering the first step.

   With probability half the first step moves right to v_k, giving us $T_k = 1$.

   With probability half the first step moves left left to v_{k-2}, giving us $T_k = 1 + X + Y$ where X has the same distribution as T_{k-1} and Y has the same distribution as T_k. We get

   $$E[T_k] = \frac{1}{2} \times 1 + \frac{1}{2} \times \Big(1 + E[X] + E[Y]\Big) = \frac{1}{2} + \frac{1}{2}\Big(1 + [2(k-1) - 1] + E[T_k]\Big).$$

   This gives $2E[T_k] = 1 + \Big(1 + [2(k-1) - 1] + E[T_k]\Big)$, or

   $$E[T_k] = 2k - 1 = k^2 - (k-1)^2.$$

4. `Problem 1.34, Stochastic Processes, Ross`. Let X_1 and X_2 be independent nonnegative continuous random variables. Let $\lambda_1(t)$ be the failure rate function of X_1. Similarly, let $\lambda_2(t)$ the failure rate function of X_2. Show

$$P\{X_1 < X_2 | \min(X_1, X_2) = t\} = \frac{\lambda_1(t)}{\lambda_1(t) + \lambda_2(t)}.$$

`Answer.` Think of X_1 as the lifetime of a device with

$$P\{X_1 \in (t, t + \mathrm{d}t) | X_1 > t\} = \lambda_1(t)\mathrm{d}t.$$

Interpret $P\{X_1 < X_2 | \min(X_1, X_2) = t\}$ as the conditional probability that device 1 fails in the time interval $(t, t+\mathrm{d}t)$ conditional on both devices working at time t but only one device working at time $t + \mathrm{d}t$. We can do this since the probability of both devices failing in an infinitesimal time interval is of the order $(\mathrm{d}t)^2$, and will not have any impact in the presence of terms of order $\mathrm{d}t$. We get

$$\begin{aligned}
& P\{X_1 < X_2 | \min(X_1, X_2) = t\} \\
= \; & P\{X_1 \in (t, t + \mathrm{d}t) \text{ and } X_2 \notin (t, t + \mathrm{d}t) | X_1 \in (t, t + \mathrm{d}t) \text{ or } X_2 \in (t, t + \mathrm{d}t)\} \\
= \; & \big[\lambda_1(t)\mathrm{d}t\big(1 - \lambda_2(t)\mathrm{d}t\big)\big] / \big[\lambda_1(t)\mathrm{d}t + \lambda_2(t)\mathrm{d}t - \big(\lambda_1(t)\mathrm{d}t \times \lambda_2(t)\mathrm{d}t\big)\big] \\
= \; & \lambda_1(t)\mathrm{d}t / \big[\lambda_1(t)\mathrm{d}t + \lambda_1(t)\mathrm{d}t\big] = \lambda_1(t) / \big[\lambda_1(t) + \lambda_2(t)\big].
\end{aligned}$$

5 Week four:Poisson processes

`Monday 18 October to Thursday 21 October 2021 (Tuesday holiday)`

1. `Tutorial on Monday 18 October`. Mijanur to do Section 1.9, Stochastic Processes, Ross.

2. `Fundamental ideas`. If we omit measure theory technicalities and focus on heuristic ideas, then, given an absolutely continuous (meaning, "spread out") inhomogeneous intensity measure $\lambda(\cdot)$ on $[0, t]$, the corresponding Poisson point process on $[0, t]$ is described by the following fundamental ideas (we write $\lambda(\mathrm{d}x)$ as shorthand for $\lambda((x, x + \mathrm{d}x))$).

   - A point is born in the infinitesimal set $\mathrm{d}x \subset [0, t]$ with an infinitesimal probability $\lambda(\mathrm{d}x)$.
   - Non overlapping infinitesimal sets give birth to points independently of each other.
   - For an infinitesimal quantity $\mathrm{d}x$ we have $1 + \mathrm{d}x = e^{\mathrm{d}x}$. In particular, since $\lambda(\mathrm{d}x)$ is also infinitesimal, since λ is "spread out", we have

     $$1 + \lambda(\mathrm{d}x) = e^{\lambda(\mathrm{d}x)}$$

     This is not true for finite y: $1 + y \neq e^y$.

     More generally, plot $1+y$ and e^y together and remember the picture. It is of fundamental importance.

- The following is also useful.
$$(1+a\mathrm{d}x)^{\frac{b}{\mathrm{d}x}} = e^{a\cdot b}.$$
If you don't want to use $\mathrm{d}x$, use ϵ and take limit $\epsilon \to 0$.

3. `Question`. Let $\lambda(\cdot)$ be a "spread out" inhomogeneous intensity measure on $[0,t]$. Consider a Poisson point process on $[0,t]$ with intensity measure $\lambda(\cdot)$. What is the probability of k points born in $[0,t]$?

   `Answer`. Partition the interval $[0,t]$ into disjoint, infinitesimal intervals, choose k of them as locations for the k points and integrate over all configurations for k locations.
$$\begin{aligned}
& P\{k \text{ points in } [0,t]\} \\
= & \prod_{\mathrm{d}x\subset[0,t],\mathrm{d}x\neq \mathrm{d}x_1,\ldots,\mathrm{d}x_k} [1-\lambda(\mathrm{d}x)] \int\cdots\int_{x_1<\ldots<x_k} [\lambda(\mathrm{d}x_1)]\times\cdots\times[\lambda(\mathrm{d}x_k)] \\
= & \prod_{\mathrm{d}x\subset[0,t],\mathrm{d}x\neq \mathrm{d}x_1,\ldots,\mathrm{d}x_k} \exp\big(-\lambda(\mathrm{d}x)\big)\frac{\int\cdots\int[\lambda(\mathrm{d}x_1)]\times\cdots\times[\lambda(\mathrm{d}x_k)]}{k!} \\
= & \exp\Big(\sum_{\mathrm{d}x\subset[0,t],\mathrm{d}x\neq \mathrm{d}x_1,\ldots,\mathrm{d}x_k} -\lambda(\mathrm{d}x)\Big)\frac{\lambda([0,t])^k}{k!} \\
= & \exp\big(-\lambda([0,t])\big)\frac{\lambda([0,t])^k}{k!}.
\end{aligned}$$

4. `Question`. Let $\lambda(\cdot)$ be on $[0,t]$. Consider a Poisson point process on $\mathbb{R}_+$ with a "spread out" inhomogeneous intensity measure $\lambda(\cdot)$ on $\mathbb{R}_+$. (Omit details about measurable sets here.) What is the probability density function for birth time of the first born point?

   `Answer`. Let $f(\cdot)$ be the probability density function on $\mathbb{R}_+$. Then $f(x)\mathrm{d}x$ is the infinitesimal probability that the first point is born in $(x, x+\mathrm{d}x)$. This can be calculated as
$$f(x)\mathrm{d}x = \prod_{\mathrm{d}u\subset[0,x]}[1-\lambda(\mathrm{d}u)]\lambda(\mathrm{d}x) = \prod_{\mathrm{d}u\subset[0,x]}\exp(-\lambda(\mathrm{d}u)\lambda(\mathrm{d}x) = \exp\big(-\lambda([0,x])\big)\lambda(\mathrm{d}x).$$

6 Week five: Poisson processes

`Monday 25 October to Thursday 28 October 2021`

6.1 `Tutorial, Monday 25 October`

Mijanur's 2^{nd} week with Section 1.9. Last part/example left for next week.

6.2 First point born on $\mathbb{R}_+$

`Question`. For the Poisson process on $\mathbb{R}_+$ with intensity measure $\lambda([0,t]) = ct$ calculate the expected time at which the first point is born.
`Answer`. Let T be the time at which the first point is born. We get
$$E[T] = \int_0^\infty t\exp(-ct)c\mathrm{d}t = \int_0^\infty t\mathrm{d}\big(-\exp(-ct)\big) = \int_0^\infty \exp(-ct)\mathrm{d}t = \frac{1}{c}.$$

6.3 k points born in a triangle

`Question`. Let $\lambda(\cdot)$ be a "spread out" (or, absolutely continuous) measure on $\mathbb{R}^2$. Since it is "spread out" a single point in $\mathbb{R}^2$ will not get strictly positive measure. Fix a set of finite measure, say, for concreteness, fix a triangle T with $\lambda(T) < \infty$. Through direct calculations calculate the probability of k points being born in the triangle T.
`Answer`. The idea is similar to what was done for a line. Break the triangle into localized infinitesimal pieces that partition the whole triangle, and on each localized infinitesimal piece a point is born independently of what happens on other (disjoint) localized infinitesimal pieces. So, if k points are born at $\mathrm{d}x_1, \ldots, \mathrm{d}x_k$ then

$$P\{k \text{ points are born in } T\} = \int_T \cdots \int_T \prod_{\mathrm{d}x \subset T, \mathrm{d}x \neq \mathrm{d}x_1, \ldots, \mathrm{d}x_k} \big(1 - \lambda(\mathrm{d}x)\big) \frac{\lambda(\mathrm{d}x_1) \cdots \lambda(\mathrm{d}x_k)}{k!}.$$

Now $1 - \lambda(\mathrm{d}x) = \exp(-\lambda(\mathrm{d}x))$ and

$$\prod_{\mathrm{d}x \subset T, \mathrm{d}x \neq \mathrm{d}x_1, \ldots, \mathrm{d}x_k} \exp(-\lambda(\mathrm{d}x)) = \exp\Big(\sum_{dx \in T \setminus \{\mathrm{d}x_1, \ldots, \mathrm{d}x_k\}} -\lambda(\mathrm{d}x)\Big) = \exp(-\lambda(T)).$$

We get

$$P\{k \text{ points are born in } T\} = \exp(-\lambda(T)) \frac{\lambda(T)^k}{k!}.$$

7 Week six: Poisson processes

`Monday 8 November to Thursday 11 November 2021`

7.1 `Tutorial, Monday 8 November`

Subrata took this tutorial (and also the previous week's tutorial in a Diwali curtailed week) and covered remaining part of Section 1.9, and Section 1.7 of Stochastic Processes by Ross.

- Markov's inequality. Here $X \geq 0$ and $a > 0$ and

  $$P\{X \geq a\} \leq E[X]/a.$$

  Convince yourself with an example that $X \geq 0$ is necessary.

- Chernoff bounds. If X is allowed to take negative values, Markov's inequality will not apply. However, $\exp(tX) > 0$ always and so qualifies for Markov's inequality. Let $a > 0$ and consider the case where $t > 0$. We get

  $$P\{X \geq a\} = P\{\exp(tX) \geq \exp(ta)\} \leq \frac{E[\exp(tX)]}{\exp(ta)}.$$

  `Question.` Do we really need $a > 0$?

- Jensen's inequality.

7.2 Time to infection of whole population

`Problem 5.14/page 288, Stochastic Processes, Ross`. A single individual is infected initially in a population of n individuals. The probability of any two individuals interacting with each other in time $\mathrm{d}t$ is $\lambda \mathrm{d}t$. Whenever an uninfected individual interacts with an infected individual, he gets infected with probability α. What is the expected time for the whole population to get infected?
`Answer`. Let T_i be the time elapsed between the i^{th} individual getting infected and the $(i+1)^{\text{th}}$ individual getting infected. We need to calculate

$$\sum_{i=1}^{n-1} E[T_i].$$

If exactly i individuals are infected, in time $\mathrm{d}t$ someone uninfected will get infected with probability $i(n-i)\alpha\lambda\mathrm{d}t$. It follows that $E[T_i] = \frac{1}{\alpha\lambda}\frac{1}{i(n-i)} = \frac{1}{\alpha\lambda n}(\frac{1}{i} + \frac{1}{n-i})$. We get

$$\sum_{i=1}^{n-1} E[T_i] = \frac{2}{\alpha\lambda n}\sum_{i=1}^{n-1}\frac{1}{i}.$$

7.3 Cars on a highway

`Problem 2.22/page 93, Stochastic Processes, Ross`. Cars enter a one-way infinite highway at Poisson rate λ. Each car entering the highway independently chooses its velocity from a common distribution F. Let $N_t\big((a,b)\big)$ be the number of cars in the interval (a,b) at time t. What is the distribution of $N_t\big((a,b)\big)$?
`Answer`.

7.4 Cars on a highway: minimizing encounters

`Definition: distribution`. For a random variable X its distribution is a probability measure on $\mathbb{R}$ such that for $A \subset \mathbb{R}$ we have

$$\mu(A) = P\{X \in A\}.$$

`Definition: median`. For a random variable X with distribution μ, a number $x \in \mathbb{R}$ is said to be median if

$$\mu\big([x,\infty)\big) \geq \frac{1}{2} \text{ and } \mu\big((-\infty,x]\big) \geq \frac{1}{2}.$$

`Problem 2.24/page 94, Stochastic Processes, Ross`. Cars enter a one-way highway of length L at Poisson rate λ. Each car entering the highway independently chooses its velocity from a common distribution F. For a car entering the highway at time t, show that as $t \to \infty$ the speed of the car that minimizes the expected number of encounters is the median of the distribution μ.
`Answer`.