Statistical Inference I: Project 1
Poisson and Renewal Process

Chi-Ning Chou
May 3, 2016

Abstract

This report is my notes on reading chapter 2 and 3 of *Stochastic Processes* by Ross [1]. The two chapters cover the Poisson process and renewal process, which are common counting process in use. In this notes, I will try to focus more on the intuitions of these two models and draw some comments about the possible reason why the researchers in the old days will come up with these brilliant ideas. Also, in the end I will share my opinion on why these two processes are important and how can they have some nice properties.

Keywords: Poisson process, renewal process

1 Introduction

Every day we are counting things. Businessmen count how many customers coming everyday, police officers count how many crimes performed every month, doctors count the number of patients getting disease etc. As a result, mathematician started to think about how to model these *counting processes*. Can we systematically and mathematically characterize the behaviors of these processes? What kinds of inference do we want to draw? How can we extend these models?

With some observations, we can find out that there are two important aspects in counting process: *time* and *number of events*, where the former is continuous and the latter is discrete. Most of the time, we fix either time or number of events and ask the behavior of the other. Namely, we may want to ask the following two questions:

- What's the number of events happened before time t?
- How long should we wait for the kth events to happen?

In our following discussions, we will focus on these two questions and extend the intuitions into more general settings.

2 Overview

In this section, we define the model for manipulating the counting ideas we mentioned in the introduction. Then, we will deduce some basic relations between time and the number of events.

2.1 Model formulation

First, let's see how to quantitatively model the counting process with respect to *time*. Please think about what kind of questions that are related to time would you want to ask in your daily counting processes?

How long should I wait for the next bus?

How long does it takes to have more than 100 customers?

No hard to see that these questions are all closely related to one issue: *waiting time*. Thus, here we formally define this concept related to time as follow.

Definition 1 (interarriving time) *Let $\{X_n : n = 1, 2, ...\}$ denotes the interarriving time, where X_n is the waiting time between the $n-1$th event and the nth event. And we denote $S_n = \sum_{i=1}^{n} X_i$ to be the total time we wait for the nth event.*

We can see that $\{X_n : n = 1, 2, ...\}$ fully determines the dynamics of the process as long as we decide the distribution of each waiting time. With this notion, we can furthermore define the counting process in another aspect: *number of events*. So, what kind of questions that are related to number of events would you want to ask in your daily counting processes?

How many typhoons have come to Taiwan this summer?

How many dumplings will be sold out in an hour?

All the questions can be summarized into how many events have happened in a certain interval of time. As a result, we formally define the concept related to number of events as follow.

Definition 2 (counting process) *Let $\{N(t) : t \geq 0\}$ denotes a counting process, where $N(t)$ is the number of events happened before time t. Formally, we have $N(t) = \sup_n\{n : S_n \leq t\}$.*

With these two definitions, we can immediately answer the question proposed in the introduction:

- What's the number of events happened before time t? **A**: $\{N(t) : t \geq 0\}$
- How long should we wait for the kth events to happen? **A**: $\{S_n : n = 1, 2, ...\}$

Now, with $\{X_n : n = 1, 2, ...\}$, $\{S_n : n = 1, 2, ...\}$ and $\{N(t) : t \geq 0\}$ we can manipulate the interactions between time and the number of events in a counting process.

Remark: Here, we haven't specify the behavior of the counting process. That is, the distribution of X_n or $N(t)$ are not determined. We will formally introduce these ideas in the next section. For now, we focus on the high-level intuition of the interaction between waiting time and the number of events.

2.2 Interaction between time and the number of events

With the above model, we may start wondering how to characterize the interaction among the two aspects of counting: time and the number of events. Let's consider Figure 1.

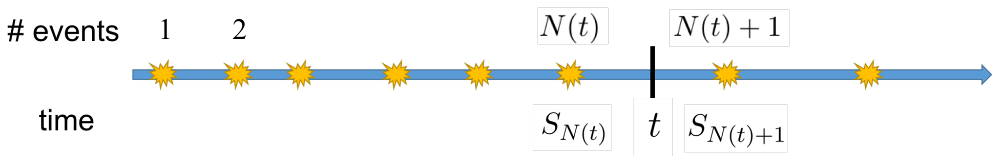

Figure 1: Interaction between time and number of events. Cut the whole process with arbitrary time n. Above the arrow is the number of events while below the line is the time.

In the following, we will see that either fix a time t or a number of events n, we can cut the whole process in to two parts with respect to both aspects. First, let's start from fixing time and consider how to cut the process into two parts concerning the number of events. From Figure 1, after cutting the process with time t, the number of events will be divided into two parts. Each part has an event that is closet to t. That is, the $N(t)$th event and $N(t) + 1$th event respectively.

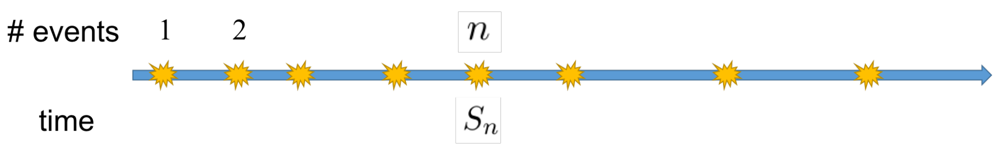

Figure 2: Interaction between time and number of events. Cut the whole process with arbitrary number of events.

On the other hands, consider Figure 2, if we fix the number of events to n, we can also cut the whole process into two parts. For time aspect, the two parts are the time smaller than S_n and the time not smaller than S_n.

To sum up, Figure 1 and Figure 2 present ways to cut the process into two parts in order to help us formalize the interaction between time and the number of events. Now, it's clear that once we choose a way to cut the process and characterize the two parts with different aspects, time and number of events will interact with each other in a certain way.

For example, cut the process with number of events n and consider the left part of the process.

- Time aspect: $\{S_n \geq t\}$
- Number of events aspect: $\{N(t) \leq n\}$

That is, we will have

$$\{S_n \geq t\} = \{N(t) \leq n\}$$

Similarly, we also have $\{S_n \leq t\} = \{N(t) \geq n\}$.

2.3 Two important structures

After we formalize the mathematical model for counting process and discuss the interaction among two different aspects, now we're going to examine two important structures in counting process: *stationarity* and *independence*. With or without these two structures will decide whether the model consists good or bad properties.

Stationarity Stationarity characterize the rules, distribution, and mechanism of the counting process. If a process has stationarity, we say it is identically distributed. That is, each waiting time have the identical distribution. On contrast, we say a process is nonhomogeneous when the distribution will varies over time. However, usually this kind of situation is difficult to deal with.

Independence Independence determines the correlation of the counting process. If a process is independent, then the waiting time after the occurrence of an event will be independent to what have happened before.

In Table 1 we summarize the terminologies of with or without the two important structures.

	With	Without	Intuition
Stationarity	Identical	Nonhomogeneous	Rules, distribution, mechanism
Independence	Independent	Correlated	Correlation, memory

Table 1: Two important structures: *stationarity* and *independence*.

Stationarity and independence help us simplify the model and thus have some great properties. However, we will also lose some precision once we assume the real world is independently and identically distributed. In the following, we will see what good results these two structures bring to us. For now, please keep in mind these two basic assumption in our models.

2.4 Motivation and applications

When given a counting process, we might want to know the following things:

- The average waiting time? (given n or asymptotic)
- The average number of events? (given t or asymptotic)
- The average behavior of the associated randomness on each event.

3 Poisson-like process and frequency

For a starter, we consider a special type counting processes that capture the idea of frequency: Poisson-like process. Basically, Poisson-like process model the counting process with a single but powerful assumption:

The probability of an event to occur in a tiny interval is independent to the history.

Here, we call this *fully independence* property.Intuitively, this assumption describe the concept of *frequency* by modeling the rate of occurrence in an interval. As a result, it becomes sufficient for us to model the frequency, which is dented as λ here. Moreover, we can even consider the frequency varying over time, which is called *nonhomogeneous* Poisson process. Let's start from the basic time-invariant Poisson process and see what's the good properties it has.

3.1 Fully stationary and independence: Poisson process

In the previous section, we observed that there are two important structures in our model for counting process: stationarity and independence. How do we *quantitatively* model these two concepts? First, independence is quite obvious, if we restrict the distribution to be independent to what have happened before, then both the waiting time and the number of events will have no correlation to the previous history. But, what about stationarity?

Semantically, stationarity refers to fixed rules, invariant distribution, or un-changed mechanism. But, we must ask what are the rues, distribution, and mechanism here mean? A quick answer is the distribution of waiting time. And this is exactly the idea of so called *renewal process*. However, the stationarity here only has meaning on the point where there's an event occurs. What about the other points?

As a result, the idea of *fully stationary* came out. What about considering that every point on the time interval have the same probability to have an event! Formally, we can write

$$P[\text{an event occurs in } [t, t+h]] = P[\text{an event occurs in } [0, h]]$$

for arbitrary $t > 0$ and h being a sufficiently small positive value. With this idea, we can define the Poisson process.

Definition 3 (Poisson process) *We say a counting process $\{N(t) : t \geq 0\}$ is a Poisson process if*

- *$\{N(t) : t \geq 0\}$ is identically distributed.*
- *$\{N(t) : t \geq 0\}$ is independent.*

Note that here we define the Poisson process only with these two properties. Now, we can quantitatively characterize the behavior of the model with a single parameter control the probability of an event to occur on a single point.

Definition 4 (Poisson process) *We say a counting process $\{N(t) : t \geq 0\}$ is a Poisson process with rate λ if*

1. $N(t) = 0$
2. $P[N(t+h) - N(t) = 1] = \lambda h + o(h),\ \forall t > 0$
3. $P[N(t+h) - N(t) > 1] = o(h),\ \forall t > 0$

With the above operational definition, we can easily derive the distribution of the number of events in a fixed period, the distribution of single (multiple) waiting time. We state them here as the following theorem.

Theorem 5 (Poisson process) *Let $\{N(t) : t > 0\}$ be a Poisson process with parameter λ. Then*

- $\forall t > 0,\ P[N(t) = n] = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$*, which is called Poisson distribution.*
- $\forall i \in \mathbb{N},\ t > 0,\ P[X_i = t] = \lambda e^{-\lambda t}$*, which is known as the exponential distribution with parameter λ.*
- $\forall n \in \mathbb{N},\ t > 0,\ P[S_n = t] = \frac{t^{n-1}\lambda^n e^{-\lambda t}}{\Gamma(n)}$*, which is known as the Gamma distribution with parameter $(n, 1/\lambda)$.*

Proof: The proof is basically a reformation of Definition 4. First, we observe the tiny variation of $N(t)$ then solve a differential equation to yield the distribution of $N(t)$ for a fixed t. The other two are then followed by the interaction between time and the number of events mentioned in Section 2.2. For more details, please refer to Rose [1]. ■

Intuition (Poisson process)

Poisson process enjoys the two important structures of renewal process: *independence* and *stationarity*. Most importantly, the independence in Poisson process is **everywhere**, which contributes to the memoryless property and thus the waiting time is exponential-like distributed.

3.2 Nonhomogeneous Poisson process

We have made a strong assumption on the Poisson process that the frequency is fixed over time. But, what if it does'nt? This will become more realistic, however, can we still have good analytical properties so that we can do some analysis? Luckily the answer is yes. Before we look into the nice results, let's start from the formal definition.

Definition 6 (nonhomogeneous Poisson process) *A counting process* $\{N(t) : t > 0\}$ *is said to be a nonhomogeneous Poisson process with frequency function* $\lambda(t)$ *if*

1. $N(t) = 0$
2. $P[N(t+h) - N(t) = 1] = \lambda(t)h + o(h),\ \forall t > 0$
3. $P[N(t+h) - N(t) > 1] = o(h),\ \forall t > 0$

We use function $\lambda(t)$ to capture the time-varying property of frequency.

With almost the same reasoning in the proof of Theorem 5, we can derive the distribution of a nonhomogeneous Poisson process as follow.

Theorem 7 (nonhomogeneous Poisson process) *Let* $\{N(t) : t > 0\}$ *be a Poisson process with parameter* λ*. Denote* $m(t) = \int_0^s \lambda(s)ds,\ \forall t > 0$*. Then*

- $\forall t > s > 0,\ P[N(t) - N(s) = n] = \frac{[m(t)-m(s)]^n e^{-[m(t)-m(s)]}}{n!}$*, which is called Poisson distribution.*
- $\forall i \in \mathbb{N},\ t > 0,\ P[X_i = t] = m'(t)e^{-m(t)}$*, which is known as the exponential distribution with parameter* λ*.*
- $\forall n \in \mathbb{N},\ t > 0,\ P[S_n = t] = \frac{m(t)^{n-1}m'(t)e^{-m(t)}}{\Gamma(n)}$*, which is known as the Gamma distribution with parameter* $(n, 1/\lambda)$*.*

Proof: Derive the distribution of the number of events in a fixed interval by solving the differential equation induced from Definition 6 in the similar way as dealing with the homogeneous Poisson process. Then, the other two are followed by the interaction between time and the number of events. ■

Intuitively, the nonhomogeneous property is averaged out by $m(t)$ in Theorem 7. $m(t)$ replaces the role of λt and $\lambda'(t)$ replace the role of λ. Since the fully independent property still hold, the nonhomogeneous Poisson process acts like being scaling from the homogeneous Poisson process with the frequency function $\lambda(t)$. Actually, this intuition can be stated as the following theorem.

Theorem 8 (scaling property) *Let* $\{N(t) : t > 0\}$ *be a homogeneous Poisson process with fixed frequency 1 and* $\{N^*(t) : t > 0\}$ *be a nonhomogeneous Poisson process with frequency function* $\lambda(t)$*. Assume the two are independent, then the following scaling property holds.*

$$P[N^*(t) = n] = P[N(m(t)) = n]$$

$\forall t > 0,\ n \in \mathbb{N}$ *and* $m(t) = \int_0^t \lambda(s)ds$*.*

Proof: Simply normalize the frequency in Definition 6 by scaling the size of the tiny interval. Then solve the differential equation as usual. ■

Intuition (nonhomogeneous Poisson process)

Once we hold the everywhere independence of Poisson process and relax the stationarity condition, we will get nonhomogeneous process. And surprisingly this relaxation preserves every good properties of Poisson process through the scaling property in Theorem 8.

3.3 Good and bad properties of Poisson process

In this subsection, I would like to share some interesting properties and variation of Poisson process without rigorous proof. The main focus will on the intuition and the implication, please refer to Rose [1] for more details.

3.3.1 Fix the number of events in an interval

When we fix the number of events in an interval, or in other words, we condition on the number of events in an interval, something cool will happen: The events in this interval will have the same distribution as the order statistics of uniform distributions!

Theorem 9 (conditioned arriving time distribute as uniform) *Given $N(t) = n$, for some $t > 0$. $S_1, S_2, ..., S_n$ have the same distributions as the order statistics of n uniformly distributed random variables on $(0, t)$.*

With this nice property, when we do some analysis on a Poisson process, we can divide the computation in two parts:

1. Fix $N(t) = n$ and do the calculation with the help of order statistics of uniform distributions.
2. Take expectation w.r.t. Poisson distribution of $N(t)$.

3.4 Extension of Poisson process

We can extend the setting of Poisson process while in the meantime still keeps the fully independence property. In this subsection, we introduce two variation: compound Poisson process and conditional Poisson process. The message here is that we still enjoy lots of good properties even we relax the setting as long as we preserve the fully independence property.

3.4.1 Compound Poisson process

There might be scenarios that something happens along with the Poisson event. For example, suppose the process is the occurrence of customers and V_i indicate the money each customer spends. Now, we might want to know the total number of money we earn in time t, namely $\sum_{i=1}^{N(t)} V_i$. That is, we associate a sequence of random variables, say V_i, to each event. And, we want to know what have happened on these random variables. Formally, we can have two aspects for the above intuitions. First, we can fix the number of events that we want to observe, and see what have happened in that period. On the other hands, we can fix the time and see what have happened. We called the first one *compound Poisson random variable* and the second *compound Poisson process.*

Definition 10 (compound Poisson random variable) *Let $\{X_i : i \in \mathbb{N}\}$ be a Poisson process with parameter λ, N be a nonnegative integer and V_i be a sequence of random variables follow the same distribution F. Define the compound Poisson random variable as*

$$W(t) = \sum_{i=1}^{N} X_i$$

Definition 11 (compound Poisson process) *Let $\{X_i : i \in \mathbb{N}\}$ be a Poisson process with parameter λ, $\{N(t) : t > 0\}$ be a Poisson process and V_i be a sequence of random variables follow the same distribution F. Define the compound Poisson process as*

$$W(t) = \sum_{i=1}^{N(t)} X_i$$

The important concept here is that we **separate** the discussion of the occurrence of events and what happen on each event into two parts. We model the occurrence with Poisson process and use another distribution F to model the randomness on each event. And surprisingly, with this separation, we can still enjoy some nice properties.

For example, there's a Poisson identity property for compound Poisson random variable so that we can compute the function of W easily.

Theorem 12 (compound Poisson identity) *With Definition 10, we have*

$$\mathbb{E}[Wh(W)] = \lambda \mathbb{E}[Xh(W + X)]$$

Basically, the message here is that with the help of compound Poisson process, we can separate the complicated process into two parts and model them respectively while in the meantime preserve some useful structure.

3.4.2 Conditional Poisson process

While nonhomogeneous Poisson process extend the Poisson process to have different occurrence rate at different time, conditional Poisson process generalize it to a probabilistic setting. Formally speaking, the rate here is no longer a determined value.

4 Renewal process

In Section 3, we saw many good properties of Poisson process and lots of useful variation. Those are mainly because of the fully independence property of Poisson assumption. What if we throw away this assumption? Soon, we find out that we still need independence to preserve good analytical structure. Aa result, we relax the concept of independence here up to the independence among two events. That is, the waiting time between event i and $i+1$ are independent to the waiting time of event j and $j+1$ as long as $i \neq j$. In this section, we are going to discover the good properties under this relaxation.

Formally, we define the renewal process as

Definition 13 (renewal process) *We say a counting process $\{N(t) : t > 0\}$ is a renewal process if the waiting time X_n follows a distribution F for all $n = 1, 2, \ldots$. Also, we denote $\mu = \mathbb{E}[X] = \int_0^\infty t dF(t)$ as the expectation of waiting time.*

For convenience, we denote F_n as the distribution for the nth waiting time, which is simply the n-fold convolution of F.

Before we start the formal discussion, there's an important function we need to define first: the *renewal function.*

Definition 14 (renewal function) *For a renewal process $\{N(t) : t > 0\}$, we define its renewal function as*

$$m(t) = \mathbb{E}[N(t)]$$

That is, the expectation of the number of events.

Clearly that $m(t)$ enjoys some good properties such as monotonicity. But, what we like renewal function so much is because of its relationship to other important properties and behaviors of the process. More examples and applications will be discussed in the following subsection. For now, we present a equation that connects renewal function and waiting distribution F to give you a taste.

$$m(t) = \sum_{n=1}^{\infty} F_n(t)$$

4.1 Four important theorems

Recall that in Section 2.4 we listed some motivations for playing around with counting process. Now, as we generalize the Poisson process to renewal process, we are intended to find whether we can answer those questions elegantly.

	Theorems	Results
Global behaviors	The elementary renewal theorem	$\frac{\mathbb{E}[N(t)]}{t} \xrightarrow{t\to\infty} \frac{1}{\mu}$
Local behaviors	Blackwell's theorem	$\mathbb{E}[N(t+a) - N(t)] \xrightarrow{t\to\infty} \frac{a}{\mu}$
Stopping time	Wald's equation	$\mathbb{E}[\sum_{n=1}^{N} X_n] = \mathbb{E}[N]\mathbb{E}[X]$
Compound properties	The key renewal theorem	$\lim\limits_{t\to\infty} \int_0^t h(t-x)dm(x) = \frac{1}{\mu} \lim\limits_{t\to\infty} \int_0^t h(x)dx$

Table 2: Four important theorems.

4.1.1 Global behaviors

We might want to know the average number of events in long term. Or even more, me might want to know the *expected* average time. Intuitively, we might guess that it will be equal to the reciprocal of the average waiting time μ. Indeed, the answer is yes for renewal process. This is stated in the *elementary renewal theorem*

Theorem 15 (the elementary renewal theorem) *By law of large number theorem, we have*

$$\frac{N(t)}{t} \xrightarrow{t\to\infty} \frac{1}{\mu}$$

Moreover, we can derive that

$$\frac{m(t)}{t} \xrightarrow{t\to\infty} \frac{1}{\mu}$$

Also, as an analogy of central limit theorem, we can derive that $N(t)$ will have some asymptotic normality property. But for now, we omit the details here.

4.1.2 Local behaviors

Previously, we consider the global behaviors of a renewal process. That is, the behaviors on the interval $[0, t]$. What if we want to know the local behaviors in $[t, t + a]$? Intuitively, we might guess that the number of events will be proportional to the length of the interval. However, will such good properties still hold? To our relief, the answer is again yes. Ans this is the famous *Blackwell's theorem.*

Theorem 16 (Blackwell's theorem) *If F is not lattice, $\forall a > 0$ we have*

$$m(t+a) - m(t) \xrightarrow{t\to\infty} \frac{a}{t}$$

Remark: We omit the discussion of lattice here for conciseness. In one sentence, lattice in the renewal process is analogous to the period in Markov chain, which is a periodic structure and will ruin some continuous property.

With Blackwell's theorem, we can make inference on an interval.

4.1.3 Compound property

Observe the Blackwell's theorem in Theorem 16, as we let $a \to 0$ and divide $m(t+a) - m(t)$ it by a, we will have

$$\lim_{t\to\infty} \frac{dm(t)}{dt} = \lim_{a\to 0} \lim_{t\to\infty} \frac{m(t+a) - m(t)}{a} = \frac{1}{\mu}$$

In other words,

$$dm(t) = \frac{dt}{\mu}, \text{ as } t \to \infty$$

This is a very powerful result since we have the following fact:

$$dm(t) = Pr[\text{event occurrs in } [t, t + dt]]$$

Thus, combine the above discussion, we have

$$Pr[\text{event occurrs in } [t, t + dt]] = \frac{dt}{\mu}, \text{ as } t \to \infty$$

That is, if we want to evaluate something when an event occur, the above results told us that we no need to care what actually F is, knowing the average is enough! This can be formally stated in the following beautiful theorem.

Theorem 17 (the key renewal theorem) *If F is not lattice and h is a Riemann integrable function, then*

$$\lim_{t\to\infty} \int_0^t h(t - x)dm(x) = \frac{1}{\mu} \lim_{t\to\infty} \int_0^t h(x)dx$$

Now, w cane sum up the above three limiting theorems for renewal process:

Intuition (Three limiting theorems)

Actually, the three theorems only tell us one thing:

The limiting (asymptotic) behavior of a renewal process is only related to μ

4.1.4 Stopping time

In stochastic process, we often care about when to **stop** the process. And this can be decided by some criteria with respect to what have happened. With this criteria, we can decide when to stop. However, we can see that the time we stop is actually a random variable because of the uncertainty of what have happened. As a results, it seems difficult to analyze the behaviors of a process with a stopping criteria, or stopping time.

But, in renewal process, the nice independent property allow us **separate** the considerations of stopping time and the behaviors of the process. Formally, we have the following Wald's equations.

Theorem 18 (Wald's equation) *Let N be a stopping time such that $\{N = n\}$ is independent to $X_{n+1}, ...$, we have*

$$\mathbb{E}[\sum_{n=1}^{N} X_n] = \mathbb{E}[N]\mathbb{E}[X]$$

5 Conclusion

To sum up, in this project, we consider two special counting process: Poisson process, and renewal process. The two processes are based on two important structural assumptions: *stationarity* and *independence*. Specifically,

- Poisson process has **fully stationarity** and **fully independence**. (The stationarity assumption can be relaxed into other variation)
- Renewal process has **event-wise independence**.

Because of these assumptions, both Poisson process and renewal process enjoy a bunch of great properties. Thus, we can easily analyze some limiting or asymptotic behaviors of these processes.

Other than the well behaviors of these two processes, let's go back to the high level view of counting process. In the beginning, we create counting process for counting sequences of discrete events over continuous/discrete time domain. For example, we can use Poisson process to model the dynamics of a markets with customers come in and pay. Or, we can use the renewal process to model the occurrence of a disease.

However, after we designed and analyzed these good models, we found out that there are a lot more applications for counting process. One interesting variation is the *alternating renewal process*, which models the switching of status in a system. The tricky part here is that we change our focus from the discrete event to the continuous intervals while in the meantime those great theorems are still there for us to use.

After studying these two processes, although I can't say I'm familiar with all of the theorems and examples. But, I kind of realizing that it is the stationarity and independence assumptions that make them have so many useful applications. Though real life does not usually have stationarity or independence, starting with these two assumptions provide us good intuitions to the behaviors of the process. As a results, in the future, if I encounter new applications issues or try to construct new models, I think finding the stationarity and independence is a nice starting point.

Another feedback is about modeling. From the example of alternating renewal process we can see that the inventor was really understand the structure of counting process so that he/she can come up with such ideas to switch the setting and still being a very useful tool. For me, this is really inspiring. And I think in the future study of new models or new topics, I will be more willing to spend more time on extracting the underlying motivation and intuition of a deep theory. From this kind of process, I kind of like building a theory with the inventor. I believe this will make me understand more about the whole great works.

References

1. Sheldon M Ross et al. *Stochastic processes*, volume 2. John Wiley & Sons New York, 1996.