

Math Review Notes—Stochastic Processes

Gregory Faletto

Contents

1	Stochastic Processes	3
1.1	Preliminaries	3
1.2	Poisson Processes	3
1.2.1	Poissonization Trick	9
1.2.2	Time Sampling Poisson Processes	14
1.2.3	Nonstationary Poisson Processes	15
1.2.4	Queueing Systems	20
1.3	Renewal Processes	21
1.4	ISE 620	31
1.5	Simple Random Walk	33
1.6	Martingales	34
1.7	Brownian Motion	34

Last updated February 21, 2019

1 Stochastic Processes

These notes are based on my notes from ISE 620 at USC, *Time Series and Panel Data Econometrics* (1st edition) by M. Hashem Pesaran and coursework for Economics 613: Economic and Financial Time Series I at USC, as well as notes from *Probability and Random Processes* by Grimmett and Stirzaker.

1.1 Preliminaries

Definition 1.1. A **stochastic process** is a collection of random variables $X(t), t \geq 0$ (in the continuous case) or X_1, X_2, \dots in the discrete case such that ...

Definition 1.2. A stochastic process $\{N(t), t \geq 0\}$ is said to be a **counting process** if $N(t)$ represents the total number of events that have occurred up to time t . Hence, a counting process $N(t)$ must satisfy

- (i) $N(t) \geq 0$
- (ii) $N(t)$ is integer-valued.
- (iii) If $s < t$ then $N(s) \leq N(t)$.
- (iv) For $s < t$, $N(s) - N(t)$ equals the number of events that have occurred in the interval $(s, t]$.

Definition 1.3. We say a counting process $\{N(t), t \geq 0\}$ has **independent increments** if the numbers of events that occur in disjoint time intervals are independent; that is, for all $t_0 < t_1 < \dots < t_n$, $N(t_1) - N(t_0), \dots, N(t_n) - N(t_{n-1})$ are independent.

Definition 1.4. A counting process $\{N(t), t \geq 0\}$ has **stationary increments** if the distribution of the number of events that occur in any interval of time depends only on the length of the time interval. That is, if $N(t+s) - N(s)$ has a distribution that does not depend on s (is the same for all s ; only depends on t).

1.2 Poisson Processes

Definition 1.5 (Poisson Process, Grimmett and Stirzaker definition). A **Poisson process with intensity λ** is a process $N = \{N(t) : t \geq 0\}$ taking values in $S = \{0, 1, 2, \dots\}$ such that

- (a) $N(0) = 0$; if $s < t$ then $N(s) \leq N(t)$.
- (b)
$$\Pr(N(t+h) = n+m \mid N(t) = n) = \begin{cases} \lambda h + o(h) & \text{if } m = 1, \\ o(h) & \text{if } m > 1; \\ 1 - \lambda h + o(h) & \text{if } m = 0 \end{cases}$$
- (c) If $s < t$, the number $N(t) - N(s)$ of emissions in the interval $(s, t]$ is independent of the times of emissions during $[0, s]$.

Remark. λ can be interpreted as the average or long-run frequency of the Poisson process.

Definition 1.6 (Poisson Process, Ross definition, 2.1.2 in Stochastic Processes). The counting process $\{N(t) : t \geq 0\}$ is said to be a **Poisson process with rate λ** if

- (a) $N(0) = 0$
- (b) $\{N(t) : t \geq 0\}$ has independent increments
- (c) $\Pr(N(t+h) - N(t) = 1) = \lambda h + o(h)$
- (d) $\Pr(N(t+h) - N(t) \geq 2) = o(h)$

Remark. Note that a Poisson process has stationary increments.

Lemma 1 (ISE 620 Ross Lemma 1). Let $N(t)$ be a Poisson process. For a fixed time s , let $N_s(t) = N(s+t) - N(s)$ be the difference between two Poisson processes. Then $\{N_s(t), t \geq 0\}$ is a Poisson process.

Proof. We verify the conditions in Definition 1.6.

- (a) $N_s(0) = N(s) - N(s) = 0$, QED.
- (b) $\{N_s(t) : t \geq 0\}$ has independent increments, yes.
- (c) $\Pr(N_s(t+h) - N_s(t) = 1) = \Pr(N(s+t+h) - N(s+t) = 1) = \lambda h + o(h)$, QED.
- (d) $\Pr(N_s(t+h) - N_s(t) \geq 2) = (\Pr(N(s+t+h) - N(s+t) \geq 2) = o(h)$, QED.

□

Lemma 2 (ISE 620 Ross Lemma 2). Let $N(t)$ be a Poisson process. Then $P_0(t) = \Pr(N(t) = 0) = e^{-\lambda t}$.

Proof.

$$\begin{aligned}
 P_0(t+h) &= \Pr(N(t+h) = 0) = \Pr(N(t+h) = 0, N(t) = 0) \\
 &= \Pr(N(t) = 0) \Pr(N(t+h) - N(t) = 0) \text{ (by independent increments)}
 \end{aligned}$$

Using conditions (iii) and (iv) of Definition 1.6,

$$\begin{aligned}
 P_0(t+h) &= P_0(t)(1 - \lambda h - o(h)) = P_0(t)(1 - \lambda h) + o(h) \\
 \iff \frac{P_0(t+h) - P_0(t)}{h} &= \frac{-\lambda h P_0(t)}{h} + \frac{o(h)}{h}
 \end{aligned}$$

Taking the limit as $h \rightarrow 0$, we have

$$\begin{aligned}
 P'_0(t) = -\lambda P_0(t) &\iff \frac{P'_0(t)}{P_0(t)} = -\lambda \iff \log(P_0(t)) = -\lambda t + C \\
 &\iff P_0(t) = \lambda e^{-\lambda t} \iff P_0(0) = 1
 \end{aligned}$$

□

Theorem 3. [Grimmett and Stirzaker theorem 6.8.2] Let $N(t)$ be a Poisson process with intensity λ . Then $N(t)$ has the Poisson distribution with parameter λt ; that is,

$$\Pr(N(t) = j) = \frac{(\lambda t)^j \exp(-\lambda t)}{j!}, \quad j = 0, 1, 2, \dots$$

Proof. See Grimmett and Stirzaker section 6.8.2, page 247. Ross proof:

$$\begin{aligned} \Pr(N(t) = n) &= \frac{1}{\Pr(X_{n+1} = t - s \mid X_n = s)} \int_0^t \Pr(N(t) = n \mid S_n = s) \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} ds = 1 \\ &= \Pr(X_{n+1} > t - s) = e^{-\lambda(t-s)} \\ \Pr(N(t) = n) &= \int_0^t e^{-\lambda(t-s)} \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} ds \\ &= \frac{e^{-\lambda t} \lambda^n}{(n-1)!} \int_0^t s^{n-1} ds = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \end{aligned}$$

□

Corollary 4. $N(t + s) - N(s) \sim \text{Poisson}(\lambda t)$

Proof. By Lemma 1, ?? is a Poisson process. ...

□

Definition 1.7 (Poisson Process, Ross definition, 2.1.1 in Stochastic Processes). The counting process $\{N(t) : t \geq 0\}$ is said to be a **Poisson process with rate λ** if

- (a) $N(0) = 0$
- (b) $\{N(t) : t \geq 0\}$ has independent increments
- (c) The number of events in any interval of length t is Poisson distributed with mean λt . That is, for all $s, t \geq 0$,

$$\Pr(N(t + s) - N(s) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, \dots$$

Remark. Note that it follows from condition (iii) in Definition 1.7 that a Poisson process has stationary increments and also that $\mathbb{E}(N(t)) = \lambda t$.

Definition 1.8 (Ross Stochastic Processes definition, Section 2.2). Let X_n denote the time between the $n - 1$ st and n th event in a Poisson process. The sequence $\{X_n, n \geq 1\}$ is called the **sequence of interarrival times**.

Definition 1.9 (Grimmett and Stirzaker definition). Let $N(t)$ be a Poisson process with intensity λ . Let T_0, T_1, \dots be given by

$$T_0 = 0, T_n = \inf\{t : N(t) = n\} \quad (1)$$

so that T_n is the time of the n th arrival. The **interarrival times** are the random variables X_1, X_2, \dots given by

$$X_n = T_n - T_{n-1}. \quad (2)$$

Remark. From knowledge of N , we can find the values of X_1, X_2, \dots by (1) and (2). Conversely, we can construct N from knowledge of the X_i by

$$T_n = \sum_{i=1}^n X_i, \quad N(t) = \max\{n : T_n \leq t\} \quad (3)$$

Theorem 5. (Grimmett and Stirzaker theorem 6.8.10.) Let $N(t)$ be a Poisson process with intensity λ . Let T_0, T_1, \dots be given by (1) and let X_n be given by (2). Then the random variables $\{X_n\}$ are independent, each having an exponential distribution with parameter λ .

Proof. See Grimmett and Stirzaker section 6.8.2, page 249. □

Corollary 6. Let $N(t)$ be a Poisson process with intensity λ . Let T_0, T_1, \dots be given by (1). Then $T_n \sim \text{Gamma}(n, \lambda^{-1})$.

Proof. By (3), $T_n = \sum_{i=1}^n X_i$. $X_i \sim \text{Exponential}(\lambda)$ by Theorem 5, which means $X_i \sim \text{Gamma}(1, \lambda^{-1})$. Then by Proposition ??, $T_n \sim \text{Gamma}(n, \lambda^{-1})$. □

Lemma 7 (ISE 620 Ross in-class Lemma 3, Proposition 2.2.1 in Stochastic Processes). Let $N(t)$ be a Poisson process. Let X_1, X_2, \dots be the interarrival times. Then X_1, X_2, \dots are independently and identically distributed as $\text{Exponential}(\lambda)$. Since the sum of exponential distributions is Gamma, we have $S_n = \sum_{i=1}^n X_i \sim \text{Gamma}(n, 1/\lambda)$.

Proof. See proof on page 64 (section 2.2) of *Stochastic Processes*. Follows from Lemmas 1 and 2. □

Remark. If the arrival times are IID exponential, then the process is Poisson. (Will come later in notes.)

Remark.

$$X_1 < t \iff N(t) = t$$

$$\Pr(X_2 > t \mid X_1 = s) = \Pr(N(s+t) - N(s) = 0 \mid X_1 = s) = \Pr(N(s+t) - N(s) = 0) = \Pr(N_s(t) = 0) = e^{-\lambda t}$$

Theorem 8. Suppose events occur in a Poisson process with parameter λ , $\text{PP}(\lambda)$. Let $N(t)$ be the number of events that occur by time t . Suppose each event is independent of all that has previously occurred; that is, each event is type $i = 1, \dots, r$ with probability p_i , and $\sum_{i=1}^r p_i = 1$. Let $N_i(t)$ be the number of type i events that occur by time t . $\{N_i(t)\} \sim \text{PP}(\lambda p_i)$. Moreover, these processes are independent for different i .

Proof. Verify axioms of Poisson process:

1. $N_i(0) = 0$
2. independent increments? Yes, the number of type i events in each interval is totally independent of what happened in previous intervals.
3. $\Pr(N_i(t+h) - N_i(t) = 1) = \Pr(N(t+h) - N(t) = 1, \text{ event is type } i + \dots$
probability of two or more events (and then one of them is type i) is $o(h)$.

$$= \lambda h p_i + o(h)$$

4. $\Pr(N_i(t+h) - N_i(t) \geq 2) \leq \Pr(N(t_h) - N(t) \geq 2) = o(h)$

Therefore $N_i(t)$ is a Poisson process with rate λp_i . We could have also proven this by looking at the inter-arrival times and showing that they are i.i.d. exponential.

□

Proof: Alternative, stronger statement.

$$\Pr(N_1(t) = n_1, N_2(t) = n_2, \dots, N_r(t) = n_r)$$

we want to show that these random variables are independent. Let $n = \sum_{i=1}^r n_i$. Then

$$= \Pr(N_1(t) = n_1, \dots, N_r(t) = n_r \mid N(t) = n) \Pr(N(t) = n)$$

But this is a Multinomial distribution. So we have

$$= \frac{n!}{n_1! \dots n_r!} \prod_{i=1}^r p_i^{n_i} \cdot e^{-\lambda t} \frac{(\lambda t)^{\sum_i n_i}}{n!}$$

$$\prod_{i=1}^r e^{-\lambda t p_i} \frac{(\lambda t p_i)^{n_i}}{n_i!}$$

And these are just Poisson probabilities.

Then using Proposition 9, we have independence (???)

□

Proposition 9 (offhand claim Ross made). If $\Pr(X = n, Y = m) = g(n)h(m)$ (that is, if probability can be broken into products of functions) then variables are independent.

Proof.

$$\Pr(X = n) = \sum_m g(n)h(m) = g(n)C_h$$

$$\Pr(Y = m) = h(m) \sum_n g(n) = h(m) C_g$$

$$1 = \sum_n \sum_m g(n) h(m) = C_g C_h$$

So this is true as long as $C_g C_h = 1$ which it does.

□

Example 1.1. Suppose in particular $\lambda = 10$ so we expect 10 people to arrive every hour, either men or women. What is the expected number of women to arrive given that 7 men arrived? (5—women and men's arrivals are independent.)

Example 1.2 (Coupon collecting problem, see also probability notes.) r types of coupons collected with probability p_1, \dots, p_r . Let N be the number of coupons you collect until you have a complete set. What is $\mathbb{E}(N)$?

Solution. Let $m(\mathcal{S})$ be the mean number of draws required to obtain at least one coupon of type i for each $i \in \mathcal{S}$. Note that $m(\emptyset) = 0$, and

$$\begin{aligned} m(\mathcal{S}) &= 1 + \sum_{i \in \mathcal{S}} p_i m(\mathcal{S} \setminus i) + \sum_{i \notin \mathcal{S}} p_i m(\mathcal{S}) \\ \implies m(\mathcal{S}) &= \frac{1 + \sum_{i \in \mathcal{S}} p_i m(\mathcal{S} \setminus i)}{\sum_{i \notin \mathcal{S}} p_i} \end{aligned}$$

But unless r is small this is going to be hard to compute, so this doesn't really work.

Solution. New idea: Let N_i be the number of draws required to obtain types $1, \dots, i$. We want N_r . Note that $\mathbb{E}(N_1) = 1/p_1$. We have $N_{i+1} = N_i + A$ where A is the additional time required.

$$\mathbb{E}(N_{i+1}) = \mathbb{E}(N_i) + \mathbb{E}(A)$$

if there is a type $i + 1$ in the original group, then $\mathbb{E}(A) = 0$. If not, note that

$$\mathbb{E}(A) = \begin{cases} 1/(p_i + 1) & \Pr(\{\text{a type } i + 1 \text{ coupon has not already been collected}\}) \\ 0 & \Pr(\{\text{a type } i + 1 \text{ coupon has already been collected}\}) \end{cases}$$

$$\implies \mathbb{E}(A) = \frac{1}{p_i + 1} \cdot \Pr(\{\text{a type } i + 1 \text{ coupon has not already been collected}\})$$

$$\implies \mathbb{E}(A) = \frac{1}{p_i + 1} \cdot \Pr(\{\text{type } i + 1 \text{ is the last of types } 1, \dots, i + 1 \text{ to be collected}\})$$

$$= \frac{1}{p_i + 1} \cdot \frac{\sum_{j=1}^{i+1} p_j}{p_{i+1}}$$

this didn't work either.

Solution.

Let N_i be the number to get type i . So $N_i \sim \text{Geometric}(p_i)$. Then $N = \max\{N_i\}$. So

$$\Pr(N \leq k) = \Pr(N_1 \leq k, \dots, N_r \leq k)$$

but you can't multiply out probabilities because N_i is not independent.

1.2.1 Poissonization Trick

Suppose we collect coupons at times distributed according to a Poisson process with rate $\lambda = 1$. Let event type i be the event of collecting a type i coupon. Let T_i be the time until collecting a type i coupon. Note that by Theorem 8 and Lemma 7, $T_i \sim \text{Exponential}(\lambda \cdot p_i) = \text{Exponential}(p_i)$. Then the time to collect at least one of every type is $T = \max_i\{T_i\}$. And by Theorem 8, the T_i are all independent. Return to the coupon collecting problem:

Solution. Recall the layer cake formulation for expected value:

$$\mathbb{E}(T) = \int_0^\infty \Pr(T > t) dt$$

We have (using independence of the T_i)

$$\Pr(T > t) = 1 - \Pr(T \leq t) = 1 - \Pr(T_1 \leq t, \dots, T_r \leq t) = 1 - \prod_{i=1}^r \Pr(T_i \leq t) = 1 - \prod_{i=1}^r (1 - e^{-p_i t})$$

$$\Rightarrow \boxed{\mathbb{E}(T) = \int_0^\infty \left(1 - \prod_{i=1}^r (1 - e^{-p_i t})\right) dt}$$

$$= \sum_i \frac{1}{p_i} - \sum_{i < j} \frac{1}{p_i + p_j} + \sum_{i, j, k} \frac{1}{[p_i + p_j + p_k]} - \dots + \frac{(-1)^{r+1}}{p_1 + \dots + p_r}$$

But we want $\mathbb{E}(N)$. Recall that T is the time of the N th event. So

$$T = \sum_{i=1}^N X_i$$

where X_1, X_2, \dots are the interarrival times of a Poisson process with rate $\lambda = 1$. But N is independent of X_1, X_2, \dots (the type of coupons you get has nothing to do with the time between coupons). So T is a compound random variable. Therefore

$$\mathbb{E}(T) = \mathbb{E}(N)\mathbb{E}(X) = \mathbb{E}(N) \cdot \frac{1}{\lambda} = \mathbb{E}(N)$$

Therefore we have $\mathbb{E}(N) = \int_0^\infty \left(1 - \prod_{i=1}^r (1 - e^{-p_i t})\right) dt$.

Example 1.3. A type is a *singleton* if after you get a complete set, you have only one object of that type. Suppose we collect coupons until we have at least one of every type, and that all coupons are equally likely to be collected on each draw. What is the expected number of singletons when you stop?

Solution. Let X be the number of singletons when you stop. Let I_j be an indicator variable for type j being a singleton. Then

$$\mathbb{E}(X) = \mathbb{E}\left[\sum_{j=1}^r I_j\right] = \sum_{j=1}^r \mathbb{E}(I_j) = \sum_{j=1}^r \Pr(\{j \text{ is a singleton}\})$$

$$\vdots$$

$$= \frac{1}{r} \sum_{i=1}^r \frac{1}{r-i+1}$$

$$\vdots$$

Let T_i be the time you get your first type i coupon. The probability we want is that at the time when you have at least one of every type except j , you have either 0 or 1 coupons of type j . That is, if $S_2^{(j)}$ is the time the second card of type j is selected, we must have $S_2^{(j)} > \max_{i \neq j} \{T_i\}$. So we seek $\Pr(S_2^{(j)} > \max_{i \neq j} \{T_i\})$. Note that $S_2^{(j)} \sim \text{Gamma}(2, p_j)$ by Corollary 6, so

$$f_{S_2^{(j)}}(s) = \frac{1}{p_j^2} s^{2-1} e^{-p_j s} = p_j e^{-p_j s} p_j \cdot s.$$

So we have

$$\begin{aligned} \Pr(I_j = 1) &= \Pr(S_2^{(j)} > \max_{i \neq j} \{T_i\}) = \int_0^\infty \Pr(\max_{i \neq j} \{T_i\} < S_2^{(j)} \mid S_2^{(j)} = s) \cdot f_{S_2^{(j)}}(s) ds \\ &= \int_0^\infty \prod_{i \neq j} (1 - e^{-p_i s}) \cdot p_j e^{-p_j s} p_j \cdot s \cdot ds \end{aligned}$$

which yields

$$\mathbb{E}(X) = \sum_{j=1}^r p_j^2 \int_0^\infty e^{-p_j s} \cdot s \cdot \prod_{i \neq j} (1 - e^{-p_i s}) \cdot ds$$

Remark. Per Example 1.4, suppose people arrive at a bus stop according to a Poisson process with rate λ . A bus arrives at (fixed) time T . Let W be the sum of the waiting times for everyone at the bus stop. Let S_j be the arrival time of the j th person. (Note that if X_i are the interarrival times then $S_j = \sum_{i=1}^j X_i$.) The number of people who get on the bus is $N(t)$, the Poisson counting process. So

$$W = \sum_{i=1}^{N(T)} (T - S_i) = N(T)T - \sum_{i=1}^{N(T)} S_i$$

Proposition 10. In a Poisson process, if we know that one event has occurred by time t , then the distribution of times of the event is uniform between 0 and t . That is,

$$S_1 < x \mid N(t) = 1 \sim U(0, t)$$

Proof.

$$\begin{aligned} \Pr(S_1 < x \mid N(t) = 1) &= \frac{\Pr(S_1 < x \cap N(t) = 1)}{\Pr(N(t) = 1)} = \frac{\Pr(N(x) = 1 \cap N(t) - N(x) = 0)}{\exp(-\lambda t)\lambda t} \\ &= \frac{\Pr(N(x) = 1) \Pr(N(t) - N(x) = 0)}{\exp(-\lambda t)\lambda t} = \frac{\lambda x e^{-\lambda x} \cdot e^{-\lambda(t-x)}}{\exp(-\lambda t)\lambda t} = \frac{x}{t} \end{aligned}$$

which is the cdf for a random variable distributed as $U(0, t)$.

□

Remark. If $X_i \sim U(0, t)$, then $f(x) = 1/t$, $0 < x < t$, so per Proposition ??,

$$f_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n) = \frac{n!}{t^n}, \quad 0 < x < t$$

Theorem 11. In a Poisson process, if we know that n events have occurred by time t ($N(t) = n$), then the set of event times $\{S_1, \dots, S_n\}$ are distributed as a set of n i.i.d. uniform random variables between 0 and t . That is, the unordered times are distributed uniformly on the interval.

More precisely: Given $N(t) = n$, S_1, \dots, S_n are distributed as order statistics of n i.i.d. $U(0, t)$ random variables.

Proof. We will examine the density function

$$f_{S_1, \dots, S_n \mid N(t)=n}(t_1, \dots, t_n), \quad 0 < t_1 < \dots < t_n < t$$

Per the above remark, we want to show that this is $n!/t^n$. Note that by Bayes' Theorem

$$f_{S_1, \dots, S_n | N(t)=n}(t_1, \dots, t_n) = \frac{f_{S_1, \dots, S_n}(t_1, \dots, t_n) \cdot \Pr(N(t) = n \mid S_1 = t_1, \dots, S_n = t_n)}{\Pr(N(t) = n)}. \quad (4)$$

Let X_1, X_2, \dots be the interarrival times. So the condition above is equivalent to $X_1 = t_1, X_2 = t_2 - t_1, \dots, X_n = t_n - t_{n-1}$. Recall that all the $\{X_i\}$ are independent. So we have

$$f_{S_1, \dots, S_n}(t_1, \dots, t_n) = f_{X_1, \dots, X_n}(t_1, t_2 - t_1, \dots, t_n - t_{n-1})$$

Also, we can interpret $\Pr(N(t) = n \mid S_1 = t_1, \dots, S_n = t_n)$ as the probability that there are 0 arrivals between times t_n and t ; that is,

$$\Pr(N(t) = n \mid S_1 = t_1, \dots, S_n = t_n) = e^{-\lambda(t-t_n)}$$

So we can write (4) as

$$\begin{aligned} f_{S_1, \dots, S_n | N(t)=n}(t_1, \dots, t_n) &= \frac{f_{X_1, \dots, X_n}(t_2 - t_1, \dots, t_n - t_{n-1}) \cdot e^{-\lambda(t-t_n)}}{e^{-\lambda t}(\lambda t)^n / n!} \\ &= \frac{\lambda e^{-\lambda t_1} \lambda e^{-\lambda(t_2-t_1)} \dots \lambda e^{-\lambda(t_n-t_{n-1})} \cdot e^{-\lambda(t-t_n)}}{e^{-\lambda t}(\lambda t)^n / n!} \\ &= \frac{e^{-\lambda t} \lambda^n}{e^{-\lambda t}(\lambda t)^n / n!} = \frac{n!}{t^n} \end{aligned}$$

□

Example 1.4. Suppose people arrive at a bus stop according to a Poisson process with rate λ . A bus arrives at (fixed) time T . What is the expected value of W , the sum of the waiting times for everyone at the bus stop?

Solution. Let S_j be the arrival time of the j th person. (Note that if X_i are the interarrival times then $S_j = \sum_{i=1}^j X_i$.) The number of people who get on the bus is $N(t)$, the Poisson counting process. So

$$\begin{aligned} W &= \sum_{j=1}^{N(T)} (T - S_j) = N(T)T - \sum_{j=1}^{N(T)} S_j \\ \mathbb{E}(W) &= \mathbb{E}(N(T)T) - \mathbb{E}\left(\sum_{j=1}^{N(T)} S_j\right) = \lambda T^2 - \mathbb{E}\left(\sum_{j=1}^{N(T)} S_j\right) \end{aligned}$$

Note that by Theorem 11, if $\{U_i\}$ are i.i.d. uniform random variables on $[0, t]$,

$$\mathbb{E}\left(\sum_{j=1}^{N(T)} S_j \mid N(T) = n\right) = \mathbb{E}\left(\sum_{j=1}^n S_j \mid N(T) = n\right) = \mathbb{E}\left(\sum_{j=1}^n U_i\right);$$

that is, if we know that n arrivals have occurred by time T , then the (unordered) arrival times are distributed uniformly on $[0, T]$. But

$$\mathbb{E}\left(\sum_{j=1}^n U_i\right) = \sum_{j=1}^n \mathbb{E}U_i = \frac{nT}{2}$$

Also note that because the U_i is independent from $N(T)$ (the arrival time of the j th person has nothing to do with how many people arrive by time T), $\sum_{j=1}^{N(T)} S_j$ is a compound random variable, which means

$$\mathbb{E}\left(\sum_{j=1}^{N(T)} S_j\right) = \mathbb{E}\left[\mathbb{E}\left(\sum_{j=1}^{N(T)} S_j \mid N(T) = n\right)\right] = \mathbb{E}\left(\frac{N(T)T}{2}\right) = \boxed{\frac{\lambda T^2}{2}}$$

\vdots

Note that the sum of the ordered values is equal to the sum of the unordered values, so we have

$$\mathbb{E}\left(\sum_{i=1}^n X_{(i)}\right) = \mathbb{E}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{E}(X_i) = n\mu$$

Example 1.5. Another thing: Let I_j be an indicator variable for the event $X_{(j)} < t$. We are interested in the distribution of $\sum_{j=1}^n I_j$.

Solution. Note that the I_j is not independent because if $I_1 = 0$ (so the smallest one is greater than t) then it must be the case that $I_2 = 0$ too. Also note that the sum of the ordered values less than t is the same as the sum of the unordered values less than t . So this is distributed as binomial with parameters n and $p = F(t)$.

Example 1.6. Suppose that shocks occur according to a Poisson process with rate λ . Let D_i be the damage caused by shock i , where $\{D_i\}$ are i.i.d. and independent of $\{N(t)\}$. The damages dissipate at an exponential rate α ; that is, damage of value d has value $de^{-\alpha s}$ after a time s . Damages are cumulative. What is the total damage by time t ?

Solution. We have

$$D(t) = \sum_{i=1}^{N(t)} D_i e^{-\alpha(t-S_i)}$$

so

$$\mathbb{E}(D(t) \mid N(t) = n) = \mathbb{E}\left(\sum_{i=1}^{N(t)} D_i e^{-\alpha(t-S_i)} \mid N(t) = n\right) = \sum_{i=1}^{N(t)} \mathbb{E}(D_i e^{-\alpha(t-S_i)} \mid N(t) = n)$$

$$= \sum_{i=1}^{N(t)} \mathbb{E}(D_i) \mathbb{E}(e^{-a(t-S_i)} \mid N(t) = n) = \mathbb{E}(D_i) e^{-at} \sum_{i=1}^n \mathbb{E}(e^{aS_i} \mid N(t) = n)$$

Let $\{U_i\}$ be i.i.d. random variables on $[0, t]$. Then by Theorem 11, we can write this as

$$\begin{aligned} &= \mathbb{E}(D_i) e^{-at} \sum_{i=1}^n \mathbb{E}(e^{aU_i} \mid N(t) = n) = \mathbb{E}(D_i) e^{-at} \cdot n \int_0^t \frac{e^{ax}}{t} dx = \mathbb{E}(D_i) e^{-at} \cdot \frac{n}{at} (e^{at} - 1) \\ &= \mathbb{E}(D_i) \cdot \frac{n}{at} (1 - e^{-at}) \end{aligned}$$

So finally we have

$$\mathbb{E}(D(t)) = \mathbb{E}[\mathbb{E}(D(t) \mid N(t) = n)] = \mathbb{E}\left[\mathbb{E}(D_i) \cdot \frac{N(t)}{at} (1 - e^{-at})\right] = \mathbb{E}(D_i) \cdot \frac{\mathbb{E}(N(t))}{at} (1 - e^{-at}) = \boxed{\mathbb{E}(D_i) \cdot \frac{\lambda}{a} (1 - e^{-at})}$$

1.2.2 Time Sampling Poisson Processes

Proposition 12 (Time Sampling (Proposition 5.3 in Introduction to Probability Models)).

Suppose we have events that happen as a Poisson process with rate λ . Each event is of type $1, \dots, r$ independently of what has come before. An event at time s is type i with probability $P_i(s)$. Let $N_i(t)$ be the number of type i events by time t . Then $N_1(t), \dots, N_r(t)$ are independent Poisson random variables with mean $\mathbb{E}[N_i(t)] = \lambda \int_0^t P_i(s) ds$.

Proof. I'm not sure this proof makes sense? Note that

$$\Pr(N_1(t) = n_1, \dots, N_r(t) = n_r) = \Pr(N_1(t) = n_1, \dots, N_r(t) = n_r \mid N(t) = \sum_{i=1}^r n_i) \cdot \Pr\left(N(t) = \sum_{i=1}^r n_i\right).$$

Given $N(t) = n$, the n event times are i.i.d. $U(0, t)$ by Theorem 11. Let P_i be the probability that a particular event is of type i given a time t and the fact that $N(t) = n$. Then

$$P_i = \int_0^t P_i(s) \cdot f_U(s) ds = \int_0^t \frac{P_i(s)}{t} ds$$

Since $\Pr\left(N(t) = \sum_{i=1}^r n_i\right) = e^{-\lambda t}$, we have

$$\Pr(N_1(t) = n_1, \dots, N_r(t) = n_r) = \frac{n!}{n_1! \dots n_r!} P_1^{n_1} \dots P_r^{n_r} e^{-\lambda t}$$

□

Remark. Notation for Queueing: in many cases we suppose times between successive arrivals are independent with a common distribution F . Then it is a renewal process. (If they are exponential, we get a Poisson process.) We have the following notation for a renewal process if k number of servers and G is the distribution of the service time:

$$F/G/k$$

If the distribution is exponential, we use M (for “memoryless” or “Markovian”). (E stands for Erlang, sum of i.i.d. exponentials).

Example 1.7. We have an $M/G/\infty$ queueing system. Arrivals are a Poisson process with rate λ . Fix t and let $X(t)$ be the number of people in the system at time t . Let $Y(t)$ be the number of people who have completed service at time t . Let the events be arrivals of customers, and call it a Type 1 event if the customer is still in the system at time t , and Type 2 if the customer completes service by time t (and Type 3 otherwise).

Suppose a customer arrives at time $s < t$. Then the customer is Type 1 if their service time exceeds $t - s$, which happens with probability $P_1(s) = \bar{G}(t - s)$. Similarly $P_2(s) = G(t - s)$. Let $N_1(t)$ be the number of Type I events that happen by time t and similarly for $N_2(t)$. Then by Proposition 12 we have

$$X(t) = N_1(t) \sim \text{Poisson} \left(\lambda \int_0^t \bar{G}(t - s) ds \right) = \text{Poisson} \left(\lambda \int_0^t \bar{G}(y) dy \right)$$

and

$$Y(t) = N_2(t) \sim \text{Poisson} \left(\lambda \int_0^t G(t - s) ds \right) = \text{Poisson} \left(\lambda \int_0^t G(y) dy \right)$$

which implies the independence of $X(t)$ and $Y(t)$.

1.2.3 Nonstationary Poisson Processes

Definition 1.10 (Nonstationary Poisson process.). The counting process $\{N(t), t \geq 0\}$ is said to be a **nonstationary Poisson process** (or **nonhomogeneous Poisson process**) with intensity function $\lambda(t), t \geq 0$ if

- (i) $N(0) = 0$.
- (ii) $\{N(t)\}$ has independent increments.
- (iii) $\Pr(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$.
- (iv) $\Pr(N(t+h) - N(t) \geq 2) = o(h)$.

Remark. Note the similarities to Definition 1.6.

Lemma 13. Let $\{N(t), t \geq 0\}$ be a nonstationary Poisson process. Let $N_s(t) = N(s+t) - N(s)$. Then $\{N_s(t), t \geq 0\}$ is a nonstationary Poisson Process with intensity $\lambda_s(t) = \lambda(s+t)$.

Remark. Note the similarity to Lemma 1.

Proof. Note the parts of Definition 1.10 above:

- (i) $N_s(0) = N(s) - N(s) = 0$.
- (ii) $\{N_s(t)\} = \{N(s+t) - N(s)\}$ has independent increments since $N(t)$ has independent increments.
- (iii) **Not sure about this part?** $\Pr(N_s(t+h) - N_s(t) = 1) = \Pr(N(s+t+h) - N(s) - [N(s+t) - N(s)] = 1) = \Pr(N(s+t+h) - N(s+t) = 1) = \lambda(t)h + o(h)$.
- (iv) $\Pr(N_s(t+h) - N_s(t) \geq 2) = o(h)$ by similar argument to (iii).

□

Definition 1.11. Let $m(t) = \int_0^t \lambda(s)ds$. Note that $m'(t) = \lambda(t)$. We call $m(t)$ the **mean value function**.

Remark. Note that the mean value function for $N_s(t)$ is

$$m_s(t) = \int_0^t \lambda_s(y)dy = \int_0^t \lambda(s+y)dy = \int_s^{s+t} \lambda(x)dx = m(t+s) - m(s).$$

Lemma 14. Let $N(t)$ be a nonstationary Poisson process. Then $\Pr(N(t) = 0) = e^{-m(t)}$ for any $t \geq 0$.

Remark. Note the similarity to Lemma 2.

Proof. Let $P(t) = \Pr(N(t) = 0)$. Then by independent increments

$$\Pr(N(t+h) = 0) = \Pr(N(t) = 0 \cap N(t+h) - N(t) = 0) = \Pr(N(t) = 0) \cdot \Pr(N(t+h) - N(t) = 0) \quad (5)$$

Let $P_0(t+h) = \Pr(N(t+h) = 0)$. Then using Definition 1.10 we have

$$\Pr(N(t+h) - N(t) = 0) = 1 - \Pr(N(t+h) - N(t) = 1) - \Pr(N(t+h) - N(t) \geq 2) = 1 - \lambda(t+h)h + o(h)$$

so we can write (5) as

$$\begin{aligned} P_0(t+h) &= P_0(t)[1 - \lambda(t+h)h + o(h)] \\ \iff \frac{P_0(t+h) - P_0(t)}{h} &= -\lambda(t+h)P_0(t) + \frac{o(h)}{h} \end{aligned}$$

Taking the limit as $h \rightarrow 0$ yields

$$\begin{aligned}
P_0'(t) = -\lambda(t)P_0(t) &\iff \int_0^s \frac{P_0'(t)}{P_0(t)} ds = \int_0^s -\lambda(t)dt \iff \log P_0(t) = \int_0^s -\lambda(t)dt \Big|_0^s \\
&\iff P_0(s) = e^{-m(s)}
\end{aligned}$$

□

Remark (Interarrival times). Let T_1 be the time of the first event. Note that $T_1 > t \iff N(t) = 0$. So

$$\bar{F}_{T_1}(t) = \Pr(T_1 > t) = \Pr(N(t) = 0) = e^{-m(t)}$$

which means

$$f_{T_1}(t) = \frac{d}{dt}(-\bar{F}_{T_1}(t)) = m'(t)e^{-m(t)} = \lambda(t)e^{-m(t)}.$$

Note that because $\lambda(t)$ is not constant, the interarrival times are not i.i.d.

Theorem 15. Let $N(t)$ be the counting process for a nonstationary Poisson process. Then $N(t) \sim \text{Poisson}(m(t))$.

Proof. We must show that

$$\Pr(N(t) = n) = e^{-m(t)} \frac{(m(t))^n}{n!}, \quad n = 0, 1, \dots$$

We have already shown that $\Pr(N(t) = 0) = e^{-m(t)} \frac{(m(t))^0}{0!} = e^{-m(t)}$ this is true when $n = 0$ in Lemma 14. We will show that this expression holds for $n = 1, 2, \dots$ by induction. Assume for a fixed n

$$\Pr(N(t) = n) = e^{-m(t)} \frac{(m(t))^n}{n!}.$$

We seek

$$\Pr(N(t) = n + 1) = \int_0^t \Pr(N(t) = n + 1 \mid T_1 = s) f_{T_1}(s) ds \quad (6)$$

Note that (using the property of independent increments)

$$\begin{aligned}
\Pr(N(t) = n + 1 \mid T_1 = s) &= \Pr(N(t) - N(s) = n \mid T_1 = s) = \Pr(N(t) - N(s) = n) \\
&= \Pr(N_s(t - s) = n)
\end{aligned}$$

By Lemma 13 above, $N_s(\cdot)$ is a Poisson process. So using that and the induction hypothesis, we have

$$\Pr(N(t) = n + 1 \mid T_1 = s) = e^{-m_s(t-s)} \frac{(m_s(t-s))^n}{n!} = e^{-[m(t)-m(s)]} \frac{[m(t)-m(s)]^n}{n!}.$$

Substituting this expression back in to (6) and using $f_{T_1}(t) = \lambda(t)e^{-m(t)}$, we have

$$\begin{aligned} \Pr(N(t) = n + 1) &= \int_0^t e^{-[m(t)-m(s)]} \frac{[m(t)-m(s)]^n}{n!} \cdot \lambda(s)e^{-m(s)} ds \\ &= \frac{e^{-m(t)}}{n!} \int_0^t [m(t)-m(s)]^n \cdot \lambda(s) ds \end{aligned}$$

Substituting $y = m(t) - m(s) \implies dy = -\lambda(s)ds$, we have

$$= \frac{e^{-m(t)}}{n!} \int_{m(t)}^0 -y^n dy = \frac{e^{-m(t)}}{n!} \int_0^{m(t)} y^n dy = e^{-m(t)} \frac{(m(t))^{n+1}}{(n+1)!}.$$

Therefore the result follows by induction. □

Corollary 16. $N(t+s) - N(s) \sim \text{Poisson}(m(t+s) - m(s)) = \text{Poisson}(\int_s^{s+t} \lambda(y) dy)$.

Proof. Follows almost immediately from Theorem 15, since if $N(t) \sim \text{Poisson}(m(t))$, □

Proposition 17. Suppose events occur according to a Poisson process with rate λ . $\{N(t)\}$ is the counting process. An event at time s is type 1 with probability $p(s)$. Let $N_1(t)$ be the number of type 1 events by time t . Then $\{N_1(t)\}$ is a nonhomogeneous Poisson Process with intensity $\lambda(t) = \lambda p(t)$.

Remark. Note the similarity to Proposition 12.

Proof. Note the parts of Definition 1.10 above:

- (i) $N_1(0) = 0$: yes.
- (ii) $\{N_1(t)\}$ has independent increments: yes, follows pretty much immediately.
- (iii) Since the probability of exactly one event in the interval that happens to be type 1 is $\lambda h P_1(t)$, we have $\Pr(N_1(t+h) - N_1(t) = 1) = \lambda h P_1(t) = \lambda(t)h + o(h)$.
- (iv) $\Pr(N_1(t+h) - N_1(t) \geq 2) = o(h)$ by similar argument to (iii).

□

So now we have that $N_1(t) \sim \text{Poisson}(\int_0^t \lambda p_1(s) ds)$. So if every time an event happens it is of type i with a certain probability that varies over time, the counting process is for each i a nonstationary Poisson process, and all the processes are independent. (Another way to understand the time sampling result from before.)

What if the arrival process is nonstationary and the probability of a type i event is also time varying? $\lambda(t)$, $p(t)$. It's a nonstationary PP with intensity $\lambda(t)p(t)$. (Probability in an interval is $\lambda(t)hp(t) + o(h)$).

One more result about $M/G/\infty$ processes:

Theorem 18. The **departure process** from an $M/G/\infty$ queue is a nonstationary Poisson process with intensity function $\lambda(t) =$

Proof. Let $D(t)$ be the number of departures by time t . Examine the axioms of Definition 1.10 above:

- (i) $D(0) = 0$: yes.
- (ii) $\{D(t)\}$ has independent increments: think of each arrival as an event. Types: type 1 if they depart in interval 1, 2 if they depart in 2, 3 if they depart in 3, 4 if they depart elsewhere, where 1, 2, and 3 are sequential nonoverlapping intervals. Note that for a given arrival time, the type of the event is dependent only on the service time. Since the arrival times are independent by assumption and the service times are also independent, the increments are independent.
- (iii) $\Pr(D(t+h) - D(t) = 1)$: Call an event type 1 if they depart in the interval $(t, t+h)$. Then $P_1(s) = G(t+h+s) - G(t+s) = g(t-s) \cdot h + o(h)$. So

$$D(t+h) - D(t) \sim \text{Poisson}(\lambda h \int_0^t g(t-s) ds + o(h)) = \text{Poisson}(\lambda h \int_0^t g(y) dy + o(h)) \text{Poisson}(\lambda h G(t) + o(h))$$

which means $\Pr(D(t+h) - D(t) = 1) = \lambda G(t) h e^{-\lambda G(t) h} =$ (by Taylor expansion of exp) $\lambda G(t) h + o(h)$ which is what we wanted to show.

- (iv) $\Pr(D(t+h) - D(t) \geq 2) = 1 - (\Pr(D(t+h) - D(t) = 0) - (\Pr(D(t+h) - D(t) = 1) = 1 - e^{-\lambda G(t) h} - \lambda G(t) h + o(h) =$ (by Taylor expansion of exp) $1 + \lambda G(t) h - \lambda G(t) h + o(h) = o(h)$.

□

Remark. Notice in the limit as $t \rightarrow \infty$ it converges to a Poisson process with rate λ .

Example 1.8. Poisson process, event i is associated with reward V_i . $\{X(t), t \geq 0\}$ is amount of money you get at time t .

$$X(t) = \sum_{i=1}^r V_i N_i(t) \tag{7}$$

or

$$\mathbb{E}(X(t)) = \sum V_i \mathbb{E}(N_i(t)) = \sum V_i \lambda \alpha_i t = \lambda t \sum \alpha_i v_i = \lambda t \mathbb{E}(X)$$

nice thing about this representation is $V_i N_i(t)$ are independent, so variance is sum of variances. Using $\text{Var}(N_i(t)) = \lambda \alpha_i t$,

$$\text{Var}(X(t)) = \sum_{i=1}^r V_i^2 \lambda \alpha_i t = \lambda t \sum V_i^2 \alpha_i = \lambda t \mathbb{E}(X^2)$$

so we verified the formulas we derived a different way.

Note that $N_i(t)$ is approximately Gaussian for large t (because Poisson goes to normal for large t). Therefore when t is large $X(t)$ is also approximately Gaussian.

1.2.4 Queueing Systems

Say we have a $M/G/1$ queueing process. That is, arrivals are $PP(\lambda)$, the service time has distribution G , and there is one server. Note that if no one is in line, the waiting time (length of an “idle period”) until the next arrival is exponential (since it’s a Poisson process). Consider the “busy periods,” that is, the time from when a customer arrives until the time the next idle period starts.

Let B be the length of a waiting period. Let S be the service time of the initial customer. Note that B is independent from what happened before. Let A be the additional time after the first customer is served until the next idle period so that $B = S + A$. Let $N(S)$ be the number of arrivals during S . Note that A depends on $N(S)$.

Suppose $N(S) = n$. If $n = 0$, then $A = 0$. If $n = 1$, A has the same distribution as B_1 . If $n = 2$, A has the same distribution as the sum of two B s. And so on. This tells us we can say

$$B = S + \sum_{i=1}^{N(S)} B_i$$

If we condition on $S = s$,

$$\{B \mid S = s\} = s + \sum_{i=1}^{N(s)} B_i$$

Note that $\sum_{i=1}^{N(s)} B_i$ is a compound Poisson random variable. So

$$\mathbb{E}(B \mid S = s) = s + \mathbb{E}(N(s))\mathbb{E}(B_i) = s + \mathbb{E}(B)\lambda s$$

$$\text{Var}(B \mid S = s) = \lambda s \mathbb{E}(B^2)$$

or

$$\mathbb{E}(B \mid S) = S + \mathbb{E}(N(S))\mathbb{E}(B_i) = S + \mathbb{E}(B)\lambda S$$

$$\text{Var}(B \mid S) = \lambda S \mathbb{E}(B^2)$$

So

$$\mathbb{E}(B) = \mathbb{E}[\mathbb{E}(B \mid S)] = \mathbb{E}(S) + \lambda \mathbb{E}(S) \mathbb{E}(B) \implies \boxed{\mathbb{E}(B) = \frac{\mathbb{E}(S)}{1 - \lambda \mathbb{E}(S)}}$$

Note the similarity to the sum of an infinite geometric series. Similarly, if $\lambda \mathbb{E}(S) \geq 1 \iff \mathbb{E}(S) \geq 1/\lambda$ then the expected waiting time is infinite because the arrival times under the Poisson process ($1/\lambda$ in expectation) are faster in expectation than the service times. Also,

$$\text{Var}(B) = \lambda \mathbb{E}(S) \mathbb{E}(B^2) + (1 + \lambda \mathbb{E}(B))^2 \text{Var}(S)$$

Note that $\text{Var}(B) = \mathbb{E}(B)^2$, etc., then you work out the answer (see textbook for complete answer.)

End of Poisson processes

1.3 Renewal Processes

Definition 1.12 (Stieltjes Integral).

$$\int_a^b h(x) dF(x) = \lim_{n \rightarrow \infty} \sum_{i=1}^n h(x_i) (F(x_i) - F(x_{i-1}))$$

In particular, we often suppose that $F(x)$ is the distribution function for a random variable X . Then we have (if X is continuous)

$$\int_a^b h(x) dF(x) = \lim_{n \rightarrow \infty} \sum_{i=1}^n h(x_i) (\Pr(x_{i-1} < X \leq x_i)) = \int_a^b h(x) f_X(x) dx = \mathbb{E}(h(x))$$

or if X is discrete

$$\int_a^b h(x) dF(x) = \lim_{n \rightarrow \infty} \sum_{i=1}^n h(x_i) (\Pr(x_{i-1} < X \leq x_i)) = \mathbb{E}(h(x)).$$

Generalization of Poisson processes. A counting process where the interarrival times are i.i.d. Let X_1, X_2, \dots be i.i.d. non-negative random variables with distribution function F . We require $F(0) = \Pr(X \leq 0) = \Pr(X = 0) < 1$. Let

$$\mathbb{E}(X_1) = \mu = \int_0^\infty x dF(x) = \int_0^\infty \bar{F}(t) dt$$

where $0 < \mu \leq \infty$. Let $N(t)$ be the counting process (this is the largest value of n for which the n th event has occurred at time t , so $N(t) = \max\{n : S_n \leq t\}$). Let S_n be the arrival time for the n th event. Define $S_0 = 0, S_n = \sum_{i=1}^n X_i$.

Definition 1.13 (Renewal process.). The counting process $\{N(t), t \geq 0\}$ is called a **renewal process**.

Every time an event occurs is a “renewal:” from that time on, all the arrivals are i.i.d. (Between events it is unclear what is going on until the next event happens; depends on the nature of F . If we have a Poisson process, then it is memoryless, but otherwise it is not.)

Example 1.9 (St. Petersburg Paradox). Idea: you play a game, you get the amount of money X . In order to play the game you have to pay, so what’s a fair amount to pay? Belief: $\mathbb{E}(X)$ (that’s a rational price).

Game: fair coin, flip until heads occurs. If it occurs on trial n , you win 2^n dollars. Note that

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} 2^i \cdot \left(\frac{1}{2}\right)^i = \infty.$$

Proposition 19. With probability 1, $N(t) < \infty$ for all t .

Proof.

$$\frac{S_n}{n} = \frac{X_1 + \dots + X_n}{n} \xrightarrow{a.s.} \mu > 0 \text{ as } n \rightarrow \infty$$

Because the denominator goes to ∞ and the ratio doesn’t go to 0, we must have $S_n \xrightarrow{a.s.} \infty$. Therefore $\Pr(N(t) < \infty) = 1$ for all t . □

Proposition 20. With probability 1, $N(\infty) = \lim_{t \rightarrow \infty} N(t) = \infty$. (There’s never a last renewal—there will always be another.)

Proof. Recall **Boole’s Inequality**:

$$\Pr\left(\bigcup_{n=1}^{\infty} \{A_n\}\right) \leq \sum_{n=1}^{\infty} \Pr(A_n)$$

(with equality if the events are disjoint). Then

$$\Pr(N(\infty) < \infty) = \Pr(X_n = \infty \text{ (for some } n)) = \Pr\left(\bigcup_{n=1}^{\infty} \{X_n = \infty\}\right) \leq \sum_{n=1}^{\infty} \Pr(X_n = \infty) = 0$$

□

Definition 1.14 (Notation for n -fold self-convolution). Recall the notation for convolution: if $X \sim F \perp\!\!\!\perp Y \sim G$, we have $X + Y \sim F * G$. Then

$$S_n \sim F * F * \dots * F(t).$$

Let

$$F_n(t) := F * F * \dots * F(t)$$

the n -fold convolution of F with itself. (In general, this is very hard to get a closed-form notation for except in some simple cases. More of a theoretical construct than something practical to compute.)

Proposition 21. $\Pr(N(t) = n) = F_n(t) - F_{n+1}(t)$.

Proof.

$$N(t) \geq n \iff S_n \leq t.$$

$$\implies \Pr(N(t) \geq n) = \Pr(S_n \leq t)$$

Then we have

$$\Pr(N(t) = n) = \Pr(N(t) \geq n) - \Pr(N(t) \geq n+1) = F_n(t) - F_{n+1}(t).$$

(Again, this is more of a theoretical construct than something practical to compute.)

□

Definition 1.15 (Renewal function). Let $m(t) = \mathbb{E}(N(t))$. We call $m(t)$ the **renewal function**.

Remark.

$$m(t) = \sum_{n=1}^{\infty} \Pr(N(t) \geq n) = \sum_{n=1}^{\infty} F_n(t).$$

Lemma 22. $m(t) < \infty$ for all t .

Proof. Skipped in class, not very enlightening.

□

Proposition 23 (Renewal Equation).

$$m(t) = F(t) + \int_0^t m(t-s) dF(s).$$

Proof.

$$m(t) = \mathbb{E}(N(t)) = \int_0^{\infty} \mathbb{E}(N(t) \mid X_1 = s) dF(s)$$

Note that

$$\mathbb{E}(N(t) \mid X_1 = s) = \begin{cases} 1 + m(t-s) & s \leq t \\ 0 & s > t \end{cases}$$

so we have

$$m(t) = \int_0^t (1 + m(t-s))dF(s) = F(t) + \int_0^t m(t-s)dF(s)$$

□

Remark. Remember we did a problem where $X_i \sim \text{Unif}(0, 1)$ and asked how many you have to sum until the number of greater than 1. We ended up with $1/e$. This is the smallest value of n for which the event occurred after time n (or something?)

$$N = \min\{X_1 + \dots + X_n > 1\} = N(1) + 1$$

(the first event that occurred after time 1. We showed that this was equal to e basically by solving the renewal equation. Only feasible because of uniform distribution between 0 and 1.

Remark. Note that

$$S_{N(t)} := (\text{time of the most recent event before or at time } t),$$

$$S_{N(t)+1} := (\text{time of the first event after time } t).$$

At what rate do renewals occur?

Theorem 24 (Strong Law for Renewal Processes).

$$\frac{N(t)}{t} \xrightarrow{a.s.} \frac{1}{\mu}$$

(Intuitively: since X_i are the time between renewals, the average time between renewals is $\mathbb{E}(X_i) = \mu$. So the average rate of renewals is equal to one over the average time between events.)

Proof. Note that

$$S_{N(t)} \leq t < S_{N(t)+1}$$

$$\iff \frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{S_{N(t)+1}}{N(t)}$$

As $t \rightarrow \infty$, the quantity on the left converges to the mean by the Strong Law of Large Numbers (see the proposition above). For the quantity on the right, we have

$$\frac{S_{N(t)+1}}{N(t)} = \frac{S_{N(t)+1}}{N(t)+1} \cdot \frac{N(t)+1}{N(t)} \xrightarrow{a.s.} \mu \cdot 1$$

which means

$$\frac{t}{N(t)} \xrightarrow{a.s.} \mu \iff \frac{N(t)}{t} \xrightarrow{a.s.} \frac{1}{\mu}.$$

□

Remark. In many applications: X_1, X_2, \dots are independent non-negative random variables. $X_1 \sim G$, $X_i \sim F, i \geq 2$. (That is, once the first event occurs we have a renewal process, but before then we don't.)

Definition 1.16 (Delayed renewal process). Define

$$N_d(t) := \max\{n : X_1 + \dots + X_n \leq t\}.$$

We call $\{N_d(t), t \geq 0\}$ a **delayed renewal process**.

Remark. Almost all limiting results for renewal processes apply for delayed renewal processes as well. (In the long run, that first waiting period doesn't really make a difference.) For example, consider the strong law for renewal processes (quick note: let $N(s) = -1, s < 0$):

$$N_d(t) = 1 + N(t - X_1) \iff \frac{N_d(t)}{t} = \frac{1}{t} + \frac{N(t - X_1)}{t - X_1} \frac{t - X_1}{t}$$

By the Strong Law for Renewal processes,

$$\frac{N(t - X_1)}{t - X_1} \xrightarrow{a.s.} \frac{1}{\mu_F}$$

Since $\lim_{t \rightarrow \infty} \frac{t - X_1}{t} = 1$, we have the same result as for the Strong Law for Renewal Processes.

Example 1.10. Arrivals to a single server queue according to a Poisson process with rate λ . However, they will only enter if the server is free when they arrive. Service time has distribution G . (This is often called a **loss model** or **Erlang loss model**, more generally with k servers.)

- (a) At what rate do customers enter the system?
- (b) What proportion of arrivals enter the system?

Solution.

- (a) Note that the events of customers entering are a renewal process because everything probabilistically starts all over again when a customer arrived. Specifically it's a delayed renewal process because the time of the first arrival is exponential with rate λ , but the rest of them are more complicated because there must be a service and an arrival (the sum of two random variables). So the rate of arrivals is 1 over the expected time between events, or

$$\frac{1}{\mathbb{E}(S + I)}$$

where S is the service time and I is the interarrival time for the next customer. Let $\mathbb{E}(X) = \mu_G = \int_0^\infty \overline{G}(t)dt$. Then

$$\mathbb{E}(S + I) = \mu_G + \frac{1}{\lambda}$$

so the answer is

$$\frac{1}{\mu_G + \frac{1}{\lambda}} = \boxed{\frac{\lambda}{1 + \lambda\mu_G}}.$$

(b) Rate of service arrivals divided by overall rates of arrivals:

$$\frac{\lambda}{1 + \lambda\mu_G} \bigg/ \lambda = \boxed{\frac{1}{1 + \lambda\mu_G}}.$$

Example 1.11. Suppose we have a bin with an infinite number of coins. Each bin has value p of landing on heads, and suppose the value of p for a randomly chosen coin is distributed as Uniform(0, 1). We draw coins and flip them. Every turn we can either draw a new coin or flip one of the coins we already have. If we want to maximize the proportion of coins that flip heads, what is the optimal strategy?

Solution. Consider a strategy of giving up on a coin if it comes up tails once. Every time you flip tails, the process renews. Let $N(n)$ be the number of tails in the first n flips. Then $N(n)$ is a renewal process. So we know that

$$\frac{N(n)}{n} \xrightarrow{a.s.} \frac{1}{\mathbb{E}(T)}$$

where T is the time between tails. Let P be a random variable for the probability of a coin flipping heads and note that

$$\mathbb{E}(T) = \int_0^1 \mathbb{E}(T \mid P = p)dp = \int_0^1 \frac{1}{1-p}dp = -\log(1-p) \Big|_0^1 = \infty$$

so the long-run proportion of coins that land heads is 1.

Definition 1.17 (Stopping times). Let X_1, X_2, \dots be independent random variables. We say the non-negative integer valued random variable N is a **stopping time** for X_1, X_2, \dots if the event $\{N = n\}$ is independent of X_{n+1}, X_{n+2}, \dots . (This definition is more general than saying it must depend on previous times because it could be random and not depend on anything (except the current one) or you could have a finite memory, etc.)

Theorem 25 (Wald's Equation). Suppose X_1, X_2, \dots are i.i.d. with $\mathbb{E}(X) < \infty$. Let N be a stopping time for X_1, X_2, \dots such that $\mathbb{E}(N) < \infty$. Then

$$\mathbb{E}\left(\sum_{i=1}^N X_i\right) = \mathbb{E}(N)\mathbb{E}(X).$$

Remark. Note the similarity to compound random variables. But the proof differs because N is not independent of the $\{X_i\}$.

Proof. Let I_i be an indicator variable for $I \leq N$. Then

$$\sum_{i=1}^N X_i = \sum_{i=1}^{\infty} X_i I_i \implies \mathbb{E}\left(\sum_{i=1}^N X_i\right) = \mathbb{E}\left(\sum_{i=1}^{\infty} X_i I_i\right) = \sum_{i=1}^{\infty} \mathbb{E}(X_i I_i)$$

Note that I_i depends on X_1, \dots, X_{i-1} but not on X_i (it only depends on whether you play the i th game, not the outcome of that game. You only decide whether to play the i th game based on the outcomes of the previous games.). Therefore $X_i \perp\!\!\!\perp I_i$. So we have

$$= \sum_{i=1}^{\infty} \mathbb{E}(X_i) \mathbb{E}(I_i) = \mathbb{E}(X) \sum_{i=1}^{\infty} \mathbb{E}(I_i) = \mathbb{E}(X) \sum_{i=1}^{\infty} \Pr(N \geq i) = \mathbb{E}(X) \mathbb{E}(N)$$

or we can write

$$\mathbb{E}(X) \sum_{i=1}^{\infty} \mathbb{E}(I_i) = \mathbb{E}(X) \mathbb{E}\left(\sum_{i=1}^{\infty} I_i\right) = \mathbb{E}(X) \mathbb{E}(N).$$

□

Remark. How do we justify

$$\mathbb{E}\left(\sum_{i=1}^{\infty} X_i I_i\right) = \sum_{i=1}^{\infty} \mathbb{E}(X_i I_i)?$$

Do the same thing again but replace all the X_i with $|X_i|$. Then the proof works, and Wald's Equation follows by Lebesgue's Dominated Convergence Theorem.

Example 1.12. Let

$$X_i = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{otherwise} \end{cases}$$

One stopping time could be $N_1 = \min\{n; X_1 + \dots + X_n = k\}$. Then by Wald's Equation,

$$\mathbb{E}\left(\sum_{i=1}^N X_i\right) = \mathbb{E}(N) \mathbb{E}(X) \iff k = \mathbb{E}(N)(p) \implies \mathbb{E}(N) = \frac{k}{p}.$$

Another could be $N_2 = \min\{n : X_{n-1} = X_n = 1\}$.

Example 1.13. Let

$$X_i = \begin{cases} 1 & \text{with probability } p \\ -1 & \text{with probability } 1 - p \end{cases}$$

with $p > 1/2$. Note that $\mathbb{E}(X_i) = 2p - 1 > 0$. One stopping time could be $N = \min\{n : X_1 + \dots + X_n = 10\}$. (Note that by the Strong Law of Large Numbers, with probability 1 this will eventually happen.) Then by Wald's Equation,

$$\mathbb{E}\left(\sum_{i=1}^N X_i\right) = \mathbb{E}(N)\mathbb{E}(X) \iff 10 = \mathbb{E}(N)(2p - 1) \implies \mathbb{E}(N) = \frac{10}{2p - 1}.$$

What's the mean amount of time until you're up by one dollar?

$$m = 1 + (1 - p)\mathbb{E}(\text{up } 2) = 1 + (1 - p)2m$$

Possible stopping rule: stop when you're winning money.

$$1 = \sum_{i=1}^N X_i = \mathbb{E}(N)\mathbb{E}(X) = 0$$

but it turns out Wald's equation doesn't apply because $\mathbb{E}(N) = \infty$ (the mean number of plays to get ahead is infinite).

Example 1.14. Let

$$X_i = \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2 \end{cases}$$

with $p > 1/2$. Note that $\mathbb{E}(X_i) = 2p - 1 > 0$. One stopping time could be $N = \min\{n : X_1 + \dots + X_n = 1\}$. It turns out from Markov chain theory that this will eventually happen with probability 1.

Example 1.15. Let $U_i \sim \text{Uniform}(0, 1)$, U_i i.i.d. One stopping time is $N = \min\{n : U_n > 0.8\}$. Then by Wald's Equation,

$$\mathbb{E}\left(\sum_{i=1}^N X_i\right) = \mathbb{E}(N)\mathbb{E}(X) = 10 \cdot 0 = 0$$

so the expected winnings when you stop is 0 (regardless of your stopping time, since it's a fair game).

Back to renewal theory: Let X_1, X_2, \dots be a sequence of random variables with interarrival times distributed as F . Note that $N(t) = 5 \iff X_1 + \dots + X_5 \leq t, X_1 + \dots + X_6 > t$ so this is not a stopping time. But we could stop at the first event that occurs after time t , so $N(t) + 1$ is a stopping time. Note that

$$N(t) + 1 = n \iff N(t) = n - 1, X_1 + \dots + X_{n-1} \leq t, X_1 + \dots + X_n > t.$$

(You can't say "I'll stop at the last event before t " because at the time of that event you wouldn't have been able to know it was the last event before t .)

Corollary 26.

$$\mathbb{E}\left(\sum_{i=1}^{N(t)+1} X_i\right) = \mu(m(t) + 1)$$

Theorem 27 (Elementary Renewal Theorem).

$$\frac{m(t)}{t} \xrightarrow{a.s.} \frac{1}{\mu}.$$

Proof. Note that

$$S_{N(t)+1} > t$$

We have

$$\mathbb{E}(S_{N(t)+1}) > t$$

Using Corollary 26,

$$\mu(m(t) + 1) > t \iff m(t) + 1 > \frac{t}{\mu} \iff \frac{m(t)}{t} > \frac{1}{\mu} - \frac{1}{t}$$

Then as $t \rightarrow \infty$

$$\liminf \frac{m(t)}{t} \geq \frac{1}{\mu}.$$

Assume $\Pr(X_i \leq M) = 1$ for some $M \in \mathbb{R}$. Then

$$S_{N(t)+1} < t + M \implies \mu(m(t) + 1) < t + M \implies m(t) < \frac{t}{\mu} + \frac{M}{\mu} - 1$$

$$\implies \frac{m(t)}{t} < \frac{1}{\mu} + \frac{M}{t\mu} - \frac{1}{t}$$

$$\implies \limsup \frac{m(t)}{t} \leq \frac{1}{\mu}.$$

Now consider the general case. Consider X_1, X_2, \dots . Let

$$X_i^* = \begin{cases} X_i & \text{if } X_i \leq M \\ M & \text{if } X_i > M \end{cases}$$

and let $N^*(t) = \max\{n : X_1^* + \dots + X_n^* \leq t\}$. Then since $N^*(t)$ has smaller interarrival times than $N(t)$,

$$\begin{aligned}
N(T) \leq N^*(t) &\implies \frac{\mathbb{E}(N(t))}{t} \leq \frac{\mathbb{E}(N^*(t))}{t} \\
&\implies \lim \frac{m(t)}{t} \leq \lim \frac{\mathbb{E}(N^*(t))}{t} = \frac{1}{\mathbb{E}(X_i^*)}
\end{aligned}$$

Note that (since $X_I^* = X_I$ if $X_i \leq M$)

$$\mathbb{E}(X_i^*) = \int_0^M x dF(x) + M\bar{F}(M) \rightarrow \mu + 0 = \mu \text{ as } M \rightarrow \infty$$

so

$$\lim \frac{m(t)}{t} \leq \frac{1}{\mathbb{E}(\min\{X, M\})}$$

which is true for every M . Letting $M \rightarrow \infty$, we have

$$\mathbb{E}(\min\{X, M\}) \rightarrow \mathbb{E}(X) \text{ as } M \rightarrow \infty$$

by Lebesgue's Monotone convergence theorem. ("if you have a sequence of variables $X_n \leq X_{n+1} \leq \dots$ then $\mathbb{E}(\lim_{n \rightarrow \infty} X_n) = \lim_{n \rightarrow \infty} \mathbb{E}(X_n)$ ").

□

Remark. This technique is called "coupling:" relating something you don't know to something you do.

Remark. Why doesn't the Elementary Renewal Theorem follow from the Strong Law for Renewal Processes? Consider this counterexample: Let $U \sim \text{Uniform}(0, 1)$ and let X_n be a random variable such that

$$X_n = \begin{cases} n & U < 1/n \\ 0 & U > 1/n \end{cases}$$

Of course with probability 1, $U > 0$; that is, $U = \epsilon > 0$. Note that $X_n = 0$ for all n sufficiently large ($1/n < \epsilon$). So we see that $X_n \xrightarrow{a.s.} 0$. But for all n sufficiently large

$$\mathbb{E}(X_n) = n \cdot \frac{1}{n} = 1.$$

So we have

$$\lim_{n \rightarrow \infty} X_n = 0 = \mathbb{E}(\lim_{n \rightarrow \infty} X_n) \neq \lim_{n \rightarrow \infty} \mathbb{E}(X_n) = 1$$

In summary, just because $\frac{N(t)}{t} \xrightarrow{a.s.} \frac{1}{\mu}$ doesn't mean $\frac{m(t)}{t} \xrightarrow{a.s.} \frac{1}{\mu}$.

1.4 ISE 620

Exercise 1. (“Best Prize” problem.) There are n prizes that are presented one at a time in a random order. Each prize has a defined value, and there is a defined ordering of the value of the prizes. Each time you see a prize, you only know the value of that prize relative to the prizes already seen—the value of later prizes remains unknown. Each time a prize is presented, we either accept it or reject it. You only get one prize, so once you accept it, you’re done. Once you reject a prize, you can’t get it back. Your goal is to maximize the probability of accepting the best prize. What do you do (what is the optimal policy)?

Solution. The only strategy that makes sense is to accept a prize if it is a *candidate*—that is, the best you’ve seen so far. You should never accept a prize that isn’t a candidate unless you get to the last prize. Let P_n be the probability of getting the best prize. Then we expect P_n to approach 0 as n approaches infinity. Note that if it were good to accept the k th prize if it were a candidate, then it would definitely be better to accept the $k + 1$ th candidate if it were a candidate. So a good policy (k -policy) is to let k prizes go by, then accept the first candidate to come afterward.

Let X be the position of the best prize. We will condition on the best prize being in position i . Then

$$P_k(\text{best}) = \sum_{i=1}^n P_k(\text{best} \mid X = i) \Pr(X = i) = \frac{1}{n} \sum_{i=1}^n P_k(\text{best} \mid X = i)$$

Note that because in order for you to win, none of the prizes between position k and i can be candidates (or else you would accept them and not get the best prize). So the best of the first $i - 1$ prizes must be among the first k prizes in order for you to get the best prize.

$$P_k(\text{best} \mid X = i) = \begin{cases} 0 & i \leq k \\ \Pr(\{\text{best of } 1, \dots, i - 1 \text{ is among the first } k\}) & k > k \end{cases}$$

and

$$\Pr(\{\text{best of } 1, \dots, i - 1 \text{ is among the first } k\}) = \frac{k}{i - 1}$$

so

$$P_k(\text{best}) = \frac{1}{n} \sum_{i=k+1}^n \frac{k}{i - 1} = \frac{k}{n} \sum_{j=k}^{n-1} \frac{1}{j} \approx \frac{k}{n} \int_k^{n-1} \frac{dx}{x} = \frac{k}{n} \log(x) \Big|_k^{n-1} = \frac{k}{n} \log\left(\frac{n-1}{k}\right) \approx \boxed{\frac{k}{n} \log\left(\frac{n}{k}\right)}.$$

Now we want to choose k that maximizes this quantity. Generalize to letting x equal any real number in the expression

$$f(x) = \frac{x}{n} \log\left(\frac{n}{x}\right) \implies f'(x) = \frac{x}{n} \frac{x}{x^2} \left(-\frac{n}{x^2}\right) + \log\left(\frac{n}{x}\right) \frac{1}{n}$$

Setting equal to 0 we have

$$\frac{1}{n} = \frac{1}{n} \log\left(\frac{n}{x}\right) \implies \frac{n}{x} = e \implies \boxed{x = \frac{n}{e}}$$

so the optimal strategy is to let about $1/e$ of the prizes go by, then choose the first candidate to come by afterward.

Remark. The probability of getting the best prize is then

$$f\left(\frac{n}{e}\right) = \frac{1}{e} \log(e) = \frac{1}{e}$$

regardless of n !

Exercise 2. (Ballot problem.) Suppose we have candidates A and B with votes counted in random order. A has n votes, B has m with $n > m$. What is the probability that A is always ahead in the count at every stage of the vote?

Solution. Let $\Pr(\{A \text{ is always ahead}\}) = P(n, m) = P_{n,m}$. Then

$$\begin{aligned} P_{n,m} &= \Pr(\{A \text{ is always ahead}\} \mid \{A \text{ receives the first vote}\}) \cdot \frac{n}{n+m} \\ &\quad + \Pr(\{A \text{ is always ahead}\} \mid \{B \text{ receives the first vote}\}) \cdot \frac{m}{n+m} \\ &= \Pr(\{A \text{ is always ahead}\} \mid \{A \text{ receives the first vote}\}) \cdot \frac{n}{n+m} + 0 \\ &= Q_{n-1,m} \cdot \frac{n}{n+m} \end{aligned}$$

where Q represents the probability that A is never behind (since going forward ties would be ok because A starts out ahead). We have

$$Q_{nm} = \Pr(\{A \text{ is never behind}\}) = \frac{n}{n_m} \Pr(\{A \text{ is never behind}\} \mid \{A \text{ gets first vote}\})$$

Note that $\Pr(\{A \text{ is never behind}\} \mid \{A \text{ gets first vote}\}) = Q_{n-1,m-1}$. But now things are more complicated because A could afford to be behind by 1 going forward. So this strategy is not working. Try instead to condition on who gets the last vote.

$$\begin{aligned} P_{n,m} &= \Pr(\{A \text{ is always ahead}\} \mid \{A \text{ receives the last vote}\}) \cdot \frac{n}{n+m} \\ &\quad + \Pr(\{A \text{ is always ahead}\} \mid \{B \text{ receives the last vote}\}) \cdot \frac{m}{n+m} \end{aligned}$$

$$= P_{n-1,m} \cdot \frac{n}{n+m} + P_{n,m-1} \cdot \frac{m}{n+m}$$

We can work this out recursively using the boundary conditions

$$P_{n,n} = 2, \quad P_{n,0} = 1, \quad n > 0$$

We will try to guess the answer.

$$P_{2,1} = \Pr(\{A \text{ gets the first two votes}\}) = \frac{1}{3}$$

$$P_{n,1} = \Pr(\{\text{first two votes are for } A\}) = \frac{n}{n+1} \cdot \frac{n-1}{n} = \frac{n-1}{n+1}$$

$$P_{3,2} = \frac{3}{5} \cdot \frac{2}{4} \cdot \Pr(\{A \text{ is not the last of the remaining votes}\}) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = \frac{1}{5}$$

$$P_{4,2} = \frac{4}{6} \cdot \frac{3}{5} \cdot \Pr(\{B \text{ does not get next two votes}\}) = \frac{4}{6} \cdot \frac{3}{5} \cdot \left(1 - \frac{2}{4} \cdot \frac{1}{3}\right) = \frac{4}{6} \cdot \frac{3}{5} \cdot \frac{5}{6} = \frac{1}{3}$$

$$P_{4,3} = \frac{4}{7} \cdot \frac{3}{6} \cdot \left(\Pr(\{\text{next is } A\}) \cdot \Pr(\{\text{not 3 } B \text{ votes in a row}\}) + \Pr(\{\text{next is } B\}) \cdot \Pr(\{$$

$$= \frac{4}{7} \cdot \frac{3}{6} \cdot \left(\frac{2}{5} \cdot \frac{3}{4} + \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} \right)$$

\vdots

It seems that

$$\boxed{P_{n,m} = \frac{n-m}{n+m}}$$

We can argue this solution is unique or we can argue it is true by induction.

1.5 Simple Random Walk

Definition 1.18. Let $\{X_i\}$ be i.i.d. We have

$$X_i = \begin{cases} 1 & \text{with probability } p \\ -1 & \text{with probability } 1-p \end{cases}$$

Let $S_k = \sum_{i=1}^k X_i$. Then $\{S_k\}$ is a **simple random walk**. (It is simple because the outcomes are either 1 or -1.)

1.6 Martingales

Definition. Let $\{y_t\}_{t=0}^\infty$ be a sequence of random variables, and let Ω_t denote the information set available at date t , which at least contains $\{y_t, y_{t-1}, y_{t-2}, \dots\}$. If $\mathbb{E}(y_t \mid \Omega_{t-1}) = y_{t-1}$ holds then $\{y_t\}$ is a martingale process with respect to Ω_t .

Definition. Let $\{y_t\}_{t=1}^\infty$ be a sequence of random variables, and let Ω_t denote the information set available at date t , which at least contains $\{y_t, y_{t-1}, y_{t-2}, \dots\}$. If $\mathbb{E}(y_t \mid \Omega_{t-1}) = 0$, then $\{y_t\}$ is a martingale difference process with respect to Ω_t .

1.7 Brownian Motion

Appendix B.13, Brownian motion. A standard Brownian motion $b(\cdot)$ is a continuous-time stochastic process associating each date $a \in [0, 1]$ with the scalar $b(a)$ such that

- (i) $b(0) = 0$
- (ii) For any dates $0 \leq a_1 \leq a_2 \leq \dots \leq a_k \leq 1$ the changes $[b(a_2) - b(a_1)], [b(a_3) - b(a_2)], \dots, [b(a_k) - b(a_{k-1})]$ are independent multivariate Gaussian with $b(a) - b(s) \sim \mathcal{N}(0, a - s)$.
- (iii) For any given realization, $b(a)$ is continuous in a with probability 1.

Other continuous time processes can be generated from the standard Brownian motion. For example, a Brownian motion with variance σ^2 can be obtained as

$$w(a) = \sigma b(a)$$

where $b(a)$ is a standard Brownian motion.

The continuous time process

$$\mathbf{w}(a) = \Sigma^{1/2} \mathbf{b}(a)$$

is a Brownian motion with covariance matrix Σ .

Definition 26 (Wiener process). Let $\Delta w(t)$ be the change in $w(t)$ during the time interval dt . Then $w(t)$ is said to follow a Wiener process if

$$\Delta w(t) = \epsilon_t \sqrt{dt}, \quad \epsilon_t \sim IID(0, 1)$$

and $w(t)$ denotes the value of the $w(\cdot)$ at date t . Clearly,

$$\mathbb{E}[\Delta w(t)] = 0, \text{ and } \text{Var}[\Delta w(t)] = dt$$

Theorem 28. Donsker's Theorem, Theorem 43, p.335, Section 15.6.3. Let $a \in [0, 1)$, $t \in [0, T]$, and suppose $(J - 1)/T \leq a < J/T$, $J = 1, 2, \dots, T$. Define

$$R_T(a) = \frac{1}{\sqrt{T}} s_{[Ta]}$$

where

$$s_{[Ta]} = \epsilon_1 + \epsilon_2 + \dots + \epsilon_{[Ta]}$$

$[Ta]$ denotes the largest integer part of Ta and $s_{[Ta]} = 0$ if $[Ta] = 0$. Then $R_T(a)$ weakly converges to $w(a)$, i.e.,

$$R_T(a) \rightarrow w(a)$$

where $w(a)$ is a Wiener process. Note that when $a = 1$, $R_T(1) = 1/\sqrt{T} \cdot S_{[T]} = 1/\sqrt{T} \cdot (\epsilon_1 + \epsilon_2 + \dots + \epsilon_T)$. Since ϵ_t 's are IID, by the central limit theorem, $R_T(1) \rightarrow \mathcal{N}(0, 1)$.

Similar (Theorem 2.1 in Phillips and Durlaf (1986)): Let $\{u_t\}$ be a sequence satisfying $\mathbb{E}(u_t) = 0$, $\gamma(0) = \mathbb{E}(T^{-1}S_t^2) \rightarrow \sigma^2 < \infty$ as $T \rightarrow \infty$, $\{u_t\}$ is square summable, $\sup_t \{\mathbb{E}(|u_t|^\beta)\} < \infty$ for some $2 \leq \beta < \infty$ and all t , $\gamma(h) = \mathbb{E}(T^{-1}(y_t - y_{t-h})^2) \rightarrow K_h < \infty$ as $\min\{h, T\} \rightarrow \infty$. Then $X_T(t) \Rightarrow W(t)$ as $T \rightarrow \infty$, where $W(t)$ is a Wiener process.

Theorem 29. Continuous Mapping Theorem (Theorem 44 of Pesaran in 15.6.3). Let $a \in [0, 1)$, $i \in [0, n]$, and suppose $(J - 1)/n \leq a < J/n$, $J = 1, 2, \dots, n$. Define $R_n(a) = n^{-1/2} S_{[na]}$. If $f(\cdot)$ is continuous over $[0, 1)$, then

$$f[R_n(a)] \xrightarrow{d} f[w(a)]$$