RENEWAL THEORY AND ITS APPLICATIONS

PETER NEBRES

Abstract. This article will delve into renewal theory. It will first look at what a random process is and then explain what renewal processes are. It will then describe, derive, and prove important theorems and formulas for renewal theory. Lastly, it will give different examples and applications of renewal theory.

Contents

1. An Introduction to Random and Renewal Processes

A random process X is a family of random variables $\{X_t : t \in T\}$ that maps from a state space Ω to some set S. The state space consists of the events that can occur for random variable X. For example, when you flip a fair coin, you expect $\Omega = \{\text{Heads}, \text{Tails}\}$. The set S is the set of the probabilities of the events in Ω. So our set S in the example of flipping a fair coin would be $S = \{\frac{1}{2}, \frac{1}{2}\}$.

It is important to note that X_t must be independent and identically distributed, which will from now on be denoted as *i.i.d.* If X_t is an *i.i.d* random variable, that means that given two $i, j \in T$, the outcome of X_i does not affect the outcome of X_j and for all $t \in T$, the distribution for X_t is the same. It is also important to note that $P(X_t)$ is the probability of X_t and

$$\sum_{i=-\infty}^{\infty} P(X_i) = 1.$$

This is the same as stating:

$$P(X_i < \infty) = 1$$

It is also important to realize that there are two types of random variables. There are discrete random variables and continuous random variables. The difference is intuitive based on the name of each variable.

Definition 1.1. A random variable X is **discrete** if the set $S = \{s_1, s_2, \dots s_n\}$ is countable. A random variable X is **continuous** if the set S is not countable.

Date: August 26, 2011.

Of course, probability functions are rarely just looking at one event. Typically, we want to look at what happens in scenario 1 given something happened in scenario 2. Or, to reword the question, if event B happens, what is the probability event A occurs? This is conditional probability.

Definition 1.2. If $P(B) > 0$, then the **conditional probability** that A occurs given that B occurs is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Something not necessarily obvious about the definition is that we did not specify whether event A was independent of event B. Well, the consequence is that whether or not event B occurs does not affect the probability that event A happens. We described what independence essentially meant, but now we have a nice mathematic definition.

Definition 1.3. Events A and B are **independent** if

$$P(A \cap B) = P(A) \cdot P(B)$$

An important thing to learn is what it means for each random variable to have the same distribution. Of course, this would be a lot easier if we knew what a distribution for a random variable was.

Definition 1.4. The **distribution function** of a random variable X is the function $F : R \to [0, 1]$ given by $F(x) = P(X \leq x)$.

Lemma 1.5. *The distribution function has these properties.*
i) The limit of $F(x)$ as $x \to -\infty$ is 0 and the limit of $F(x)$ as $x \to \infty$ is 1.
ii)If $x < y$, then $F(x) \leq F(y)$
iii) F is right continuous.

So what does the distribution function give you? Well, intuitively, it tells you the area of the curve up to that point. Of course, it is also important to know why this distribution function is useful. One important result of the distribution function is that you can determine $P(a < X \leq b)$ using $F(x)$. $P(a < X \leq b)$ is $F(b) - F(a)$. There are also different distributions based on the type of random variable. A discrete random variable would have a distribution function that looked like

$$F(x_0) = \sum_{x \leq x_0} P(X = x)$$

given $x, x_0 \in S$. If the discrete random variable used a discrete sum, the continuous random variable would of course use a continuous sum, which is the integral. Its distribution function looks like

$$F(x) = \int_{-\infty}^{x} f(u)du$$

$f : R \to [0, \infty)$ is the **probability density function**, where f is integrable. Though $F(x)$ is not necessarily always differentiable (the coin flipping example, for instance), if F is differentiable at u, then $F'(u) = f(u)$. $f(x)$ is very helpful when dealing with continuous random variables because we can integrate $f(x)$, as opposed to dealing with a messy $F(x)$. $f(x)$ has the following traits.

Lemma 1.6. *If X has density function f, then i)* $\int_{-\infty}^{\infty} f(x)dx = 1$,
ii)$P(X = x) = 0$ *for all* $x \in R$,
iii)$P(a \leq x \leq b) = \int_a^b f(x)dx$.

These traits are similar to the traits for the distribution function. So, as a general guide, you use $F(x)$ when dealing with discrete random variables and $f(x)$ when dealing with continuous random variables. Of course, we also want to look at distributions when the two events are not independent of each other. For that we will use a joint probability distribution and a joint density function.

Definition 1.7. The **joint probability function** of random variables X and Y is the function $F : R^2 \to [0, 1]$ given by

$$F(x, y) = P(X \leq x, Y \leq y)$$

Given what we know about probability functions, this definition makes a lot of sense. We keep everything essentially the same, but we had a new random variable Y. We should expect a similar outcome with the joint density function.

Definition 1.8. Random variables X and Y are **jointly continuous** and have a **joint (probability) density function** $f : R^2 \to [0, \infty)$ if

$$F(x, y) = \int_{v=-\infty}^{x} \int_{u=-\infty}^{y} f(u, v)dudv \text{ for each } x, y \in R$$

Of course, now that we have a formula for when they are dependent, naturally, we want to find other equations for density or distribution functions for other situations especially the sum of random variables. The density function for a sum of random variables X and Y is actually very easy to find.

Theorem 1.9. *If X and Y have joint density function f, then $X + Y$ has density function,*

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x, z - x)dx$$

This is a very nice equation for multiple reasons. One is that we only have to look at one variable x. There are no multiple integrations, which simplifies the process, but it does not make it simple. It is also nice to know that the density function of the sum is determined by the joint density function, as that is not necessarily obvious. But this is an important theorem because much of this paper will deal with a process that is the sum of random variables.

An important parameter when it comes to random processes is the expectation value.

Definition 1.10. Expected Value of a discrete random variable X_t

$$E(X_t) = \sum_{j=1}^{\infty} P(X_t = j) \cdot j$$

This is equivalent to saying:

$$E(X_t) = \sum_{j=1}^{\infty} P(X_t > j)$$

When we looked at distribution functions for discrete and continuous random variables, we had to use different notations. We cannot use a similar formula because with the continuous random variable, $P(X_t = j) = 0$. But we can use the density function to help us find the expected value. We also cannot use the summation, so again, we use the integral.

Definition 1.11. The Expectation Value of a continuous random variable

$E(X_t) = \int_{-\infty}^{\infty} xf(x)$ given $f(x)$ the distribution of X_t

Here's an example of what the expected value can show. If we consider the random process of a fair six-sided die with sides of values $1, 2, \ldots, 6$, then the expected value is $\sum_{j=1}^{6} 1/6 \cdot j = \frac{7}{2}$. $E(X_t)$ here shows the average value of the die over a long period of time. There is another important remark about the expected value that will come in handy later.

$$E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) \tag{1.12}$$

This is a very important consequence. Think about this example. We have John and Jane. Jane wants to roll a die 100 times and find the sum. John is lazier than Jane so he rolls the die only once and multiples that value by 100. What equation 1.12 states is that if this is repeated enough times, John and Jane should have similar values. This is the linearity of the expected value.

Another important aspect of the expectation value is the conditional expectation. This looks similar to the typical expected value, but we use a conditional probability instead.

Definition 1.13. Conditional Expectation Value
If we look at a discrete random variable to a range χ,

$$E(X|Y = y) = \sum_{x \in \chi} xP(X = x|Y = y)$$

We only look at the discrete random variables because problems arise when we start to deal with random variables and their conditional expected value. This definition of course follows from what we know about expected values, but why is this important? One important effect that it has is that it allows us to look at multiple events at once. It also gives us a way to deal with expected values of events that are not independent. But it also may help when we deal with random variables that are created through other random variables through basic operations like addition. Essentially, we can deal with more types of random variables. But, for our intents and purposes, it is important to note that conditional expected value may come in handy when dealing with sums of *i.i.d.* random variables.

2. Renewal Process

A renewal process looks at successive occurrences of events in terms of random variables. But mathematically speaking, what is a renewal process?

Definition 2.1. Renewal process
A renewal process $N = (N(t) : t \geq 0)$ is a process for which

$$N(t) = \max\{n : T_n \leq t\}$$

where

$$T_0 = 0, T_n = X_1 + X_2 + \ldots + X_n \text{ for } n \geq 1$$

for $i.i.d$ non-negative random variables X_i for all $i \geq 0$

It is important to say that X_j is non-negative. Now the probability distribution that we are looking at is strictly positive. That is to say:

$$\sum_{i=0}^{\infty} P(X_i) = 1. \tag{2.2}$$

This means that we don't even have to look at any number less than 0 for our distribution!

Note by equation 1.12 and the fact that we are using $i.i.d$ random variables

$$E(T_n) = E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$$

so $E(T_n) = nE(X_1)$.

But what does a renewal process look like? The easiest way to explain this would be making an example. We shall use the example with the coin flipping. We will once again set

$$X_j = \begin{cases} 0 & \text{if tails,} \\ 1 & \text{if heads.} \end{cases}$$

So what does $N(t)$ stand for in this case? Well $N(0) = 0$, which shows how many tails have been flipped. Since we are looking strictly at non-negative Z, when $a \leq |x| < b$, it means that $N(x) = N(a)$. So, in this case, $N(\frac{1}{2}) = N(\frac{42}{43}) = N(0) = 0$. $N(1)$ would equal the time it takes to flip one heads. Similarly $N(2)$ would equal the time it takes to flip two. So $N(t)$ would equal the time it takes to flip n heads.

This example shows a generalization of these renewal processes. We shall consider T_n as the 'time of the nth arrival' and X_n as the 'nth interarrival time'. We hope that our renewal process is honest, or that every possibility from $t \geq 1$ has some $N(t) \neq 0$. That is to say that we don't want $N(t)$ to be clustered around 0. Well, let's find a parameter that we can use to determine whether or not our process is honest. One of the parameters we have is the expected value. We know that $E(X_i)$ is determined by $P(X_i = j)$ and j. The cases we should consider are $E(X_i) < 0$, $E(X_i) = 0$, and $E(X_i) > 0$. Since $j \geq 0$ and $P(X_i = j) \geq 0$, we only have to consider the latter two cases. So if $E(X_i) = 0$, then $P(X_i = j) = 0$ or $j = 0$ for all i. That means that when $E(X_i) = 0$, $P(X_i < \infty) \neq 1$ or $j = 0$. So intuitively, to have an honest process, $E(X_i) > 0$, bringing us to the theorem.

Theorem 2.3. *$P(N(t) < \infty) = 1$ for all t if and only if $E(X_1) > 0$.*

The proof relies on an important observation that we can derive from the definition of a renewal process.

$$N(t) \geq n \Longleftrightarrow T_n \leq t \tag{2.4}$$

Proof. If we consider $E(X_1) = 0$, then $P(X_i = 0) = 1\ \forall i$. From the definition of a renewal process, it is implied that

$$P(N(t) = \infty) = 1 \text{ for all } t > 0$$

Now suppose that $E(X_1) > 0$. That means that there exists an $\epsilon > 0$ such that $P(X_1 > \epsilon) = \delta > 0$. Let $A_i = \{X_i > \epsilon\}$, and let $A = \limsup A_i$ be when infinitely many X_i exceed ϵ. So

$$P(A^c) = P(\bigcup_m \bigcap_{n>m} A_n^c) = \sum_m \lim_{n\to\infty}(1-\delta)^{n-m} = \sum_m 0 = 0$$

If we apply equation (2.4),

$$P(N(t) = \infty) = P(T_n \leq t \ \forall \ n) \leq P(A_c) = 0.$$

□

From Theorem 2.3, we know that $P(X_i = 0) < 1$, but we can make the stronger claim, that $P(X_i = 0) = 0$. So we only consider X_i strictly positive. We can now get many important lemmas that follow from the previous theorem and equation (2.4). We can actually find the distribution of $N(t)$ in terms of the distribution of our random variable. Let F be the distribution of X_i and let F_n be the distribution of T_n.

Lemma 2.5. $F_1 = F$ *and* $F_{k+1}(x) = \int_0^x F_k(x-y)\mathrm{d}F(y)$ *for* $k \geq 1$.

The proof of 2.5 follows from theorem 1.9 and that $F_1 = F$ and $T_{k+1} = T_k + X_{k+1}$.

Lemma 2.6. $P(N(t) = k) = F_k(t) - F_{k+1}(t)$

Proof. $\{N(t) = k\} = \{N(t) \geq k\}/\{N(t) \geq k+1\}$. Then use 2.4. □

We now have an idea of the distribution of $N(t)$. We should also look at the expected value of it.

Definition 2.7. The **renewal function** m is given by $m(t) = E(N(t))$.

It is easy to find m in terms of F_k and in terms of F.

Lemma 2.8. *Relation between Renewal Function and Distribution Functions*

In terms of F_k*:*

$$m(t) = \sum_{k=1}^{\infty} F_k(t). \tag{2.9}$$

In terms of F *(also known as the* ***renewal equation****):*

$$m(t) = F(t) + \int_0^t m(t-x)dF(x). \tag{2.10}$$

Proof. The proof of equation 2.9 requires the use of a different set of variables. Define the indicator variables

$$I_k = \begin{cases} 1 & \text{if } T_k \leq t, \\ 0 & \text{otherwise.} \end{cases}$$

Then $N(t) = \sum_{k=1}^{\infty} I_k$ and so

$$m(t) = E(\sum_{k=1}^{\infty} I_k) = \sum_{k=1}^{\infty} E(I_k) = \sum_{k=1}^{\infty} F_k(t).$$

The proof for equation 2.10 requires the conditional expectation which gives us

$$m(t) = E(N(t)) = E(E[N(t)|X_1]);$$

If $t < x$, then

$$E(N(t)|X_1 = x) = 0$$

because the first arrival occurs after time t. If $t \geq x$, then

$$E(N(t)|X_1 = x) = 1 + E(N(t - x))$$

because the process is a copy of N. So it follows that

$$\begin{aligned} m(t) = \int_0^t E(N(t)|X_1 = x)dF(x) &= \int_0^t [1 + m(t - x)]dF(x) \\ &= \int_0^t dF(x) + \int_0^t m(t - x)dF(x) \\ &= F(t) + \int_0^t m(t - x)dF(x) \end{aligned}$$

□

This is a very important consequence because it will help us when we start to look at our renewal process as we approach infinity, which is one of the most interesting parts of this branch of probability.

From the previous lemmas, we now know that

$$\sum_{k=1}^{\infty} F_k(t) = F(t) + \int_0^t m(t - x)dF(x)$$

Interestingly, if we look at bounded intervals, equation 2.9 is the unique solution to equation 2.10. We also can generalize equation 2.10 to consider rather than just our distribution function for any uniformly bounded function H. We can look at solutions μ to the **renewal-type equation**

$$\mu(t) = H(t) + \int_0^t \mu(t - x)dF(x), t \geq 0 \tag{2.11}$$

This brings us to the next theorem.

Theorem 2.12. *The function μ, given by*

$$\mu(t) = H(t) + \textstyle\int_0^t H(t - x)dm(x),$$

is a solution of the renewal-type equation. If H is bounded on finite intervals, then μ is bounded on finite intervals and is the unique solution of the renewal-type equation.

So what does the renewal-type equation and this theorem tell us about renewal processes?

Now that we covered what should happen with the renewal process, something that would be more interesting would be to look at what happens when we look as $t \to \infty$.

3. Limit Theorems in Renewal Processes

This will cover what happens asymptotically to $N(t)$ and $m(t)$. For example, we'd hope that as $t \to \infty$ that $N(t) \to \infty$ as well. If we go back to the coin flipping case, what we are saying is that we want to show that we will get infinitely many heads over an infinite amount of time. So, there are four important results. There are two for $N(t)$ and two for $m(t)$.

The first observation for $N(t)$ is a law of large numbers for renewal processes. We have $\mu = E(X_1)$ and $\mu < \infty$.

Theorem 3.1. *$\frac{1}{t}N(t) \to \frac{1}{\mu}$ as $t \to \infty$.*

Proof. First note that $T_{N(t)} \leq t \leq T_{N(t)+1}$ If $N(t) > 0$, then

$$\frac{T_{N(t)}}{N(t)} \leq \frac{t}{N(t)} \leq \frac{T_{N(t)+1}}{N(t)+1}\Big(1 + \frac{1}{N(t)}\Big)$$

Then, as we take $t \to \infty$, then, almost surely, $N(t) \to \infty$, and we get

$$\mu \leq \lim_{t\to\infty}\Big(\frac{t}{N(t)}\Big) \leq \mu$$

□

The next trait we look at deals with the Central Limit Theorem or CLT. In general, the CLT states that under a sufficiently large number of our random variables, our distribution will eventually look normal. So, given $\sigma^2 =$var(X_1) what we want is something that resembles

$$\frac{N(t) - E(N(t))}{\sqrt{\text{var}(N(t))}} \to N(0,1)$$

where $N(0,1)$ is the normal curve with mean of 0 and standard deviation of 1.

Well, we know that $E(N(t)) = m(t)$, but we will soon learn that as $t \to \infty$, $\frac{m(t)}{t} \to \frac{1}{\mu}$. So we know that $E(N(t)) = m(t) = \frac{t}{\mu}$ as $t \to \infty$. So now we have

$$\frac{N(t) - \frac{t}{\mu}}{\sqrt{\text{var}(N(t))}} \to N(0,1)$$

So what is the variance of $N(t)$? Well we know that $E(N(t)) = nE(X_1)$, so there may be a relationship with var$(N(t))$, $E(X_1)$, $E(N(t))$, or var(X_1).

There are two ways to look at this. We know that the var$(N(t))$ =var$(\sum_{i=1}^n X_i)$ so that means that var$(N(t)) = \sum_{i=1}^n \sum_{j=1}^n \text{cov}(X_i, X_j)$ where cov(X_i, X_j) is the covariance of X_i and X_j. It is possible to simplify this further. We could consider (with proof, of course) that var(X_i) =var(X_j) $\forall$ i and j, then var$(N(t)) = n^2$var(X_1). The alternative is using the fact that var$(N(t)) = E(N(t)^2) - E(N(t))^2$.

Unfortunately, none of these methods are rather helpful. Of these methods, one of them does not work and the other two are very tedious. But the theorem is important and it gives the standard deviation function for our renewal process.

Theorem 3.2. *If $\sigma^2 =var(X_1)$ satisfies $0 \leq \sigma \leq \infty$, then*

$$\frac{N(t) - t/\mu}{\sqrt{t\sigma^2/\mu^3}} \to N(0,1) \text{ as } t \to \infty$$

If we use this information, we learn that $E(N(t)^2) = E(N)\Big(\frac{var(X_1)}{E(X_1)^2} + E(N)\Big)$, which is a very convoluted answer. The proof of this uses equation 2.6 and requires a basic knowledge of the normal curve.

The next theorem describes a trait of $m(t)$ as $t \to \infty$. It was necessary in the previous part when looking at $E(N(t))$. This theorem is called the **elementary renewal theorem**.

Theorem 3.3. *$\frac{m(t)}{t} \to \frac{1}{\mu}$ as $t \to \infty$*

In order to prove Theorem 3.3, we need another lemma. This lemma revolves around a random variable M which maps to the set $\{1, 2, \ldots\}$. We call M a *stopping time* with respect to X_i if, for all $m \geq 1$, the event $\{M \leq m\}$ belongs to the σ-field of events generated by $X_1, X_2, \ldots, X_m$. For example, $N(t) + 1$ is a stopping time, while $N(t)$ is not. A good example of what constitutes a stopping time is this. Say you were going to a friend's house and he tells you to turn right at the last stop light, or the stop light before the last one. The problem is that you don't know when the last or second to last stop light is. But if he told you to turn after the second railroad crossing, you would know when to turn. That is because the railroad crossing is a stopping time. You can determine when to stop, something you cannot do when you look at the stop lights. Now that we know what a stopping time is, here is the important lemma called **Wald's equation** [1].

Lemma 3.4. *Let* $X_1, X_2, \ldots$ *be* i.i.d *random variables with finite mean, and let* M *be a stopping time with respect to the* X_i *satisfying* $E(M) < \infty$. *Then*

$$E\Big(\sum_{i=1}^{M} X_i\Big) = E(X_1)E(M)$$

Proof. Note that

$$\sum_{i=1}^{M} X_i = \sum_{i=1}^{\infty} X_i I_{\{M \geq i\}}$$

Therefore,

$$E\Big(\sum_{i=1}^{M} X_i\Big) = \sum_{i=1}^{\infty} E(X_i I_{\{M \geq i\}}) = \sum_{i=1}^{\infty} E(X_i) P(M \geq i)$$

The final sum is

$$E(X_1) \sum_{i=1}^{\infty} P(M \geq i) = E(X_1)E(M)$$

□

Now that we have Wald's equation, we can successfully prove the elementary renewal theorem.

Proof. We know that $t < T_{N(t)+1}$ and if we take the expected values and use Wald's equation, we get

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

As $t \to \infty$, we get

$$\liminf_{t \to \infty} \frac{1}{t} m(t) \geq \frac{1}{\mu} \tag{3.5}$$

We then have to truncate X_i. It may seem easier to bound $m(t)$ using a similar method as we used in the proof for Theorem 3.1

$$t \geq E(T_{N(t)+1}) = E(T_{N(t)+1} - X_{N(t)+1}) = \mu[m(t) + 1] - E(X_{N(t)+1}) \tag{3.6}$$

[1] We know that it's important because it's named after someone.

but since $X_{N(t)+1}$ is determined by $N(t)$, it cannot be that $E(X_{N(t)+1}) = m(t)$. So we truncate X_i at some a and we get

$$X_j^a = \begin{cases} X_j & \text{if } X_j < a, \\ a & \text{if } X_j \geq a. \end{cases}$$

We look at N^a with interarrival times X_j^a and we apply equation 3.6 with $\mu^a = E(X_j^a) \leq a$ and get

$$t \geq \mu^a[E(N^a(t)) + 1] - a \tag{3.7}$$

We also know that for all j, $X_j^a \leq X_j$ and therefore for all t, $N^a(t) \geq N(t)$. So

$$E(N^a(t)) \geq E(N(t)) = m(t)$$

and then with some technical algebra, equation 3.7 is equivalent to

$$\frac{m(t)}{t} \leq \frac{1}{\mu^a} + \frac{a - \mu^a}{\mu^a * t}$$

We then take $t \to \infty$ and get

$$\limsup_{t\to\infty} \frac{1}{t} * m(t) \leq \frac{1}{\mu^a}$$

We take $a \to \infty$ and note $\mu^a \to \mu$ and get

$$\limsup_{t\to\infty} \frac{1}{t} * m(t) \leq \frac{1}{\mu}$$

□

This theorem is very, very nice. It essentially states that as long as t is sufficiently large, we know how it acts. Consider this scenario. We could have a renewal process that does not have its first interarrival moment until time 100. That would be similar to flipping a coin 100 times and not getting a head until the 100th flip. This renewal process will look different than other renewal processes with a small enough t. But what this theorem states is that we know how $m(t)$ will act as $t \to \infty$.

The next theorem involving $m(t)$ requires an extra definition.

Definition 3.8. Call a random variable X and its distribution F_X **arithmetic with span** λ (> 0) if X takes values in the set $\{m\lambda : m = 0 \pm 1, \ldots\}$ with probability 1, and λ is maximal with this property.

Theorem 3.9. ***Renewal Theorem*** *If X_1 is not arithmetic then*

$$m(t + h) - m(t) \to \frac{h}{\mu} \text{ as } t \to \infty \text{ for all } h. \tag{3.10}$$

If X_1 is arithmetic with span λ, then 3.10 holds whenever h is a multiple of λ.

It is also useful to generalize this in a theorem known as the **key renewal theorem**.

Theorem 3.11. ***Key Renewal Theorem*** *If $g : [0, \infty) \to [0, \infty)$ is such that:*
i) $g(t) \geq 0$ *for all* t,
ii) $\int_0^\infty g(t)dt < \infty$,
iii) g *is a non-increasing function,*
then

$$\int_0^t g(t - x)dm(x) \to \frac{1}{\mu} \int_0^\infty g(x)dx \text{ as } t \to \infty$$

The proof for these theorems are very technical and require a certain level of mathematics that is beyond what we will cover [2]. But the consequences of these theorems are very important. The key renewal theorem is an integral version of the renewal theorem. These theorems give us the solution to the renewal equation that we covered in the last section. Certain applications of these theorems include allowing us to find the limiting mean excess. The most important consequence of these theorems although is the fact that we now know how these processes work at large enough t. It is also very important to realize that we have a general understand of how all of these processes work as $t \to \infty$, regardless of how these renewal processes start.

4. Examples of Renewal Processes

There are many examples of renewal processes. One of the most often cited ones is the lightbulb example. The light bulb example deals with the lifespan of a light bulb. We have one light bulb in a room and we turn it on. We keep the light bulb on until it runs out and then we switch to the next light bulb. We make the assumption that switching the light bulb takes absolutely no time.

First, we should make sure that our random variable is necessarily *i.i.d.* We can assume the life span of one light bulb does not certainly affect the life span of another. We can also safely assume that the life spans are identically distributed if we have sufficiently many light bulbs. Obviously if we have only two light bulbs and they come from the same box, it is pretty obvious that they probably were not randomly chosen and in all likelihood, both will have similar life spans. But if we have essentially an infinite number of light bulbs, we do not run into this problem. So we do have independent, identically distributed light bulbs

So what does T_n stand for? Well, let's assume that you change the light bulb every four hours or $t = 4k$ just to get an initial idea. So $T_1 = 0$, because you are still on the first light bulb. But $T_4 = 1$, because you had to change the light bulb once. Following that, $T_8 = 2$, $T_12 = 3$, and so on. It also follows that $T_4.8 = 1$, $T_9 = 2$, and so on. So, it seems that T_n in this scenario is how many times the light bulb has changed at time n. But in order to determine the other qualities of our renewal process, we need to give X a distribution, which we will do later in this section.

Another interesting consequence of renewal processes is the **inspection paradox**. So we have our renewal process $N(t)$ with $t \geq 0$ with interarrival distribution F and with X_i as the ith interarrival time. So the paradox states that the distribution is stochastically larger than that of X_1; that is,

$$P\{X_{N(t)+1} > x\} \geq 1 - F(x)$$

In layman's terms, inspection paradox is this. Say we are waiting for a bus. There are 5 buses over the hour, so the mean wait time between buses is 12 minutes. At any random time in a 12 minute interval, you'd expect to wait 6 minutes, but the inspection paradox has it that the expected time of waiting is more than 6 minutes. That is to say that we are more likely to choose a period where there is a longer gap between buses[3].

[2]The proof of 3.9 was done by Torgny Lindvall in 1977.

[3]Just the beginning of the proof of Murphy's Law

We can prove this easily if we look at the age of the renewal process $A(t)$. $A(t)$ is equal to the time at t since the last renewal, assuming a renewal occurred at $t = 0$. If $A(t) = s$, then

$$P\{X_{N(t)+1} > x | A(t) = s\} = P\{X_1 > x | X_1 > s\} \geq P\{X_1 > x\}$$

Another interesting example of a renewal process is the **Poisson process**. A Poisson process is the only renewal process that has the Markov property. The Markov property is

$$P(X_n = s | X_0 = x_0, X_1 = X_1, \ldots, X_{n-1} = x_{n-1}) = P(X_n = s | X_{n-1} = x_{n-1})$$

We define a Poisson process as

Definition 4.1. A **Poisson process** is a process $N = \{N(t) : t \geq 0\}$ which satisfies the following properties:
i) The number of changes in nonoverlapping intervals are independent for all intervals,
ii) The probability of exactly one change in a sufficiently small interval $h \equiv \frac{1}{n}$ is $P = vh \equiv \frac{v}{n}$, where v is the probability of one change and n is the number of trials.
iii) The probability of two or more changes in a sufficiently small interval h is essentially 0.

It is also important to note that $N(t)$ has a Poisson distribution, which is

$$f(k; \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

where λ is the expected number of occurrences (mean) and k is the number of occurrences. We will show that the Poisson process is a renewal process. [4] Intuitively, the Poisson process is just a renewal process with exponential interarrival times. The distribution is $F(X) = 1 - e^{\lambda x}$ and $E(X) = \frac{1}{\lambda}$. So since we know that the Poisson distribution is a renewal process, let us set that as our distribution for X_i in the example at the beginning of this section. The most important and interesting question is what $m(t)$ is equal to. We already know that the mean is λ, so we should expect something similar. By equation 2.10, we know that

$$m(t) = \sum_{k=1}^{\infty} \int_0^t \frac{\lambda(\lambda s)^{k-1} e^{-ks}}{(k-1)!} ds = \int_0^t \lambda ds = \lambda t$$

We can also get this if we note that $N(t)$ has the Poisson distribution with parameter λt. So that is the distribution and expected value of our lightbulb example given that it has a Poisson distribution.

I will leave you with one more interesting paradox. This paradox will require me to define certain parameters.

Definition 4.2. :
i) The **excess lifetime** at t is $B(t) = T_{N(t)+1} - t$.
ii) The **current lifetime** or **age** at t is $A(t) = t - T_{N(t)}$.
iii) The **total lifetime** at t is $C(t) = B(t) + A(t) = X_{N(t)+1}$.

The paradox is the **waiting time paradox**. It states that if N is a Poisson process with parameter λ, then $E(B(t))$ could be one of two things.
I) Since N is a Markov chain, the distribution of $B(t)$ does not depend on arrivals

[4] It is also the only renewal process with the Markov property.

prior to time t. So $B(t)$ has the same mean as $B(0) = X_1$, and so $E(B(t)) = \lambda^{-1}$.
II) If t is large, then on average it is near the midpoint of the interarrival interval I_t which contains it. So $E(B(t)) = \frac{1}{2\lambda}$.
So this basically states that $\frac{1}{2} = 1$, which we know is not true. We actually know from the proof of the elementary renewal theorem that the reasoning for part II is not correct. In fact, $X_{N(t)+1}$ is special. Longer interval times have a higher chance of having t in the interior than smaller intervals. Furthermore, $E(B(t)) = E(X_{N(t)+1}) = \frac{(2-e^{-\lambda t})}{\lambda}$. And lastly, in this process, $B(t)$ and $C(t)$ is independent for all t, something true only for this renewal process.

Acknowledgments. It is a pleasure to thank my mentor Zachary Madden, who helped me with finding resources, editing my paper, and understanding the mathematics at hand in the article. I would like to thank my parents for giving me a place to stay, reasonable transportation costs, and delicious food. I would like to thank Ben Riffer-Reinert for helping me with LaTeX and getting me both the program and a guide that made this paper possible. I would like to thank Prof. Peter May and his two grad students who helped make REU a wonderful learning experience. I would lastly like to thank all of the professors who helped broaden the world of mathematics and were willing to give their time over the summer to students of mathematics like me who were excited to learn something new.

References

[1] Ben Parker Damn Lies. http://plus.maths.org/content/damn-lies
[2] Chapter 3 Renewal Theory. $http : //www.postech.ac.kr/class/ie272/ie666_temp/Renewal.pdf$
[3] Geoffrey R. Grimmett and David R. Stirzaker. Probability and Random Processes. Oxford University Press. 2001.
[4] Sheldon M. Ross. The Inspection Paradox. Probability in the Engineering and Informational Sciences, 17, 2003. 47-51.
[5] Karl Sigman. IEOR 6711: Notes on the Poisson Process 2009.