Poisson processes (and mixture distributions)

James W. Daniel
Austin Actuarial Seminars
www.actuarialseminars.com

June 26, 2008

Foreword

This document covers the material on Poisson processes needed for Exams MLC/3L of the Society of Actuaries (SoA) and Casualty Actuarial Society (CAS). It concentrates on explaining the ideas and stating the important facts rather than deriving the theory. In order to conserve space, rather than containing problems it instead lists problems for practice that can be downloaded from the SoA website starting at
http://www.soa.org/education/resources/edu-multiple-choice-essay-examinations.aspx
and from the CAS website starting at
http://www.casact.org/admissions/studytools/ .

Chapter 1

Poisson processes

1.1 What's a Poisson process?

Let's make our way towards a definition of a Poisson process. First of all, a Poisson process N is a stochastic process—that is, a collection of random variables $N(t)$ for each t in some specified set. More specifically, Poisson processes are counting processes: for each $t > 0$ they count the number of "events" that happen between time 0 and time t. What kind of "events"? It depends on the application. You might want to count the number of insurance claims filed by a particular driver, or the number of callers phoning in to a help line, or the number of people retiring from a particular employer, and so on. Whatever you might mean by an "event",

$$\boxed{N(t) \text{ denotes the number of events that occur after time 0 up through and including time } t > 0.}$$

$N(0)$ is taken to equal 0—no events can have occurred before you start counting.

Since the times at which events occur are assumed to be random, $N(t)$ is a random variable for each value of t. Note that N itself is called a random ***process***, distinguishing it from the random ***variable*** $N(t)$ at each value of $t > 0$.

To understand counting processes, you need to understand the meaning and probability behavior of the ***increment*** $N(t+h) - N(t)$ from time t to time $t+h$, where $h > 0$ and of course $t \geq 0$. Since $N(t+h)$ equals the random number of events up through $t+h$ and $N(t)$ equals the random number up through t, the increment is simply the random number of events occurring strictly after time t and up through and including time $t+h$. Note that $N(t)$ itself can be viewed as an increment, namely from time 0 to time t, since $N(t) = N(t) - N(0)$.

A fundamental property of Poisson processes is that increments on non-overlapping time intervals are independent of one another as random variables—stated intuitively, knowing something about the number of events in one interval gives you no information about the number in a non-overlapping interval. But notice the important modifier "non-overlapping". While the increments $N(5.2)-N(3.1)$ and $N(2.7)-N(1.6)$ ***are*** independent, $N(5.2)-N(3.1)$ and $N(4.3)-N(3.7)$ are ***not*** independent since the time intervals overlap—knowing that, say, six events occur between times 3.7 and 4.3 would tell you that at least six occur between times 3.1 and 5.2. Similarly, the increments $N(5.2) - N(3.1)$ and $N(3.8) - N(1.5)$ are ***not*** independent since the time intervals overlap. Note, however, that $N(5.2) - N(3.1)$ and $N(3.1) - N(2.4)$ ***are*** independent since the intervals do not overlap—the first increment starts counting events *strictly after* time 3.1, while the second stops counting events *at* time 3.1; that is, touching at an endpoint is "OK"—time intervals that touch only at an endpoint *do* have independent increments.

The final ingredient for a definition of a Poisson process is the rate at which events are, on average, occurring. The ***rate (or intensity) function*** λ gives the rate as $\lambda(t)$ at time t. Note that the rate can vary with time, just as the speed of a car—the rate at which it is covering distance—can vary with time. You get the total distance covered by a car during a time interval by multiplying the rate by the length of the time interval, provided the rate is a constant; if the rate varies, you integrate the rate function over the interval, which gives the same result as the simpler formula when the rate is constant. Similarly, you integrate a Poisson process's rate function over an interval to get the average number of events in that interval.

It's almost time for the definition. Since the definition of a Poisson process refers to a Poisson ***random variable*** with mean Λ, I first want to remind you about Poisson random variables.

Definition & Properties 1.1 (Poisson random variable) *A Poisson random variable M with mean Λ is a random variable with the following properties:*

1. *its only possible values are the non-negative integers 0, 1, 2, 3, ...; and*
2. $\Pr[M = k] = e^{-\Lambda}\Lambda^k/k!$.

   From these properties it follows that
3. $\mathrm{E}[M] = \mathrm{Var}[M] = \Lambda$.

The following definition formalizes the preceding discussion of Poisson processes.

KEY $\Rightarrow$ **Definition & Properties 1.2 (Poisson process)** *A Poisson process N with rate function λ has the following properties:*

1. *N is a counting process—$N(0) = 0$, and for $t > 0$ $N(t)$ is non-decreasing and takes on only non-negative integer values 0, 1, 2, 3, ... (and so can be viewed as the random number of events of interest occurring after time 0 and by time t);*
2. *N has independent increments—any set of increments $N(t_j + h_j) - N(t_j)$ for $j = 1, 2, \ldots, n$ is independent for each positive integer n, provided that the time intervals $(t_j, t_j + h_j]$ are non-overlapping (touching at an endpoint is OK); and*
3. *for all $t \geq 0$ and $h > 0$, the increment $N(t + h) - N(t)$ is a Poisson random variable with mean $\Lambda = \int_t^{t+h} \lambda(z)\,dz$.*

   Some additional notation and terminology follow:
4. *the function m defined by $m(t) = \int_0^t \lambda(z)\,dz$ is called the* mean-value function, *since* $\mathrm{E}[N(t)] = m(t)$*; in actuarial literature, m is often called the* operational time*; and*
5. *if the rate function λ is in fact a constant, then N is called a* homogeneous *Poisson process.*

Exams! $\Rightarrow$ In common usage and on the actuarial exams, "Poisson process" has usually meant "homogeneous Poisson process", while "non-homogeneous Poisson process" has been used to indicate a rate function that is not a constant. But the terminology shouldn't confuse you, since you can always just look to see whether λ is a constant. You also see statements like "Events occur at the Poisson rate 3 per hour"; this is shorthand for "Events are occurring according to a Poisson process with constant rate function $\lambda = 3$ per hour".

Let's look at an example of how the properties of a Poisson process are used, especially that of independent increments.

Example 1.3 Suppose that N is a Poisson process with rate function given by $\lambda(t) = 2t$. Also suppose that you've already observed 100 events by time 1.2—that is, that $N(1.2) = 100$. [By the way, this is a surprisingly large number, since $N(1.2) = N(1.2) - N(0)$ is a Poisson random variable with mean just $\Lambda = \int_0^{1.2} 2z\,dz = 1.44$.] Taking this observation of $N(1.2)$ into account, you want to understand the probability behavior of $N(2.6)$. You might (momentarily) wonder whether you should expect events to continue occurring surprisingly often. But then you remember that Poisson processes have independent increments, so knowing what happened from time 0 to time 1.2 should have no impact on what happens from time 1.2 to time 2.6.

More precisely, write $N(2.6)$ as

$$N(2.6) = [N(2.6) - N(1.2)] + [N(1.2) - N(0)].$$

Given the observation that $N(1.2) = 100$, you know that the second term in square brackets [ ] equals 100. Because of independent increments, given the observation the first term in square brackets is simply a Poisson random variable—call it M—with mean $\Lambda = \int_{1.2}^{2.6} 2z\,dz = 5.32$. Thus, given the observation, the distribution of $N(2.6)$ is given by

$$[N(2.6) \,|\, N(1.2) = 100] \sim M + 100,$$

with M a Poisson random variable with mean 5.32 (and with $\sim$ meaning "has the same probability distribution as").

This makes it easy to translate probability questions about $N(2.6)$, given $N(1.2) = 100$, to questions about M. For example, for the mean

$$\mathrm{E}[N(2.6) \,|\, N(1.2) = 100] = \mathrm{E}[M + 100] = \mathrm{E}[M] + 100 = 5.32 + 100 = 105.32,$$

and for the variance

$$\mathrm{Var}[N(2.6) \,|\, N(1.2) = 100] = \mathrm{Var}[M + 100] = \mathrm{Var}[M] = 5.32;$$

the equality of $\mathrm{Var}[M + 100]$ and $\mathrm{Var}[M]$ in the above followed from the fact that adding a constant to a random variable shifts all its random values by the same amount but not how those values vary among themselves. You compute probabilities just as easily; for example,

$$\Pr[N(2.6) = 104 \,|\, N(1.2) = 100] = \Pr[M + 100 = 104] = \Pr[M = 4] = \frac{e^{-5.32}(5.32)^4}{4!} = 0.16330,$$

where I calculated the final probability using property 2) in Definition & Properties 1.1 on Poisson random variables. ¶

1.1.1 Some important time intervals

Two crucial time intervals are the interval from time 0 until event number n—the n^{th} ***event time***—and the time from event number $j - 1$ to event number j—the j^{th} ***interevent time***.

Definition 1.4 (time intervals) *The random variable equal to the time of the n^{th} event—the n^{th}* ***event time****—is denoted by T_n; for convenience, T_0 is defined as $T_0 = 0$. The random variable equal to the time from event number $j - 1$ to event number j—the j^{th}* ***interevent time****—is denoted by V_j; for convenience, V_1 is defined as $V_1 = T_1$.*

It should be obvious from the definition that

$$V_j = T_j - T_{j-1}, \qquad \text{and } T_n = V_1 + V_2 + \cdots + V_n. \tag{1.5}$$

On occasion, you need to compute probabilities involving event times T_n. I recommend that you avoid doing so directly by instead translating the problem to one involving $N(t)$, a Poisson random variable for which probabilities are easy to compute. I'll show you what I mean in an example.

Example 1.6 As in Example 1.3, suppose that N is a Poisson process with rate function given by $\lambda(t) = 2t$. Suppose you're asked to compute $\Pr[T_3 > 1.2]$. Think about the event "$T_3 > 1.2$": that just says that the third event occurs after time 1.2. What does that say about $N(1.2)$, the number of events that ***have*** occurred by time 1.2? Since the third event occurs *after* time 1.2, it must be that at most two events have occurred by time 1.2—that is, that $N(1.2) \leq 2$. Conversely, if $N(1.2) \leq 2$ then at most two events have occurred by time 1.2 and so the time T_3 of the third event must be later: $T_3 > 1.2$. That is, the two events "$T_3 > 1.2$" and "$N(1.2) \leq 2$" are equivalent and so have equal probabilities. But, as you saw in Example 1.3, $N(1.2)$ is a Poisson random variable (M, say) with mean 1.44. So

$$\begin{aligned}\Pr[T_3 > 1.2] &= \Pr[N(1.2) \leq 2] = \Pr[M \leq 2] = \Pr[M = 0] + \Pr[M = 1] + \Pr[M = 2] \\ &= e^{-1.44} + e^{-1.44}\frac{1.44^1}{1!} + e^{-1.44}\frac{1.44^2}{2!} = 0.82375. \quad \P\end{aligned}$$

Example 1.7 Suppose that N is a Poisson process with rate function λ, and that at a particular time x you are interested in the random variable $T(x)$ defined as the time from x until the next event occurs—that is, if $N(x) = n$, then $T(x) = T_{n+1} - x$. Since the event $T(x) > t$ is equivalent to the event $N(x+t) - N(x) = 0$, it's easy to describe the probability behavior of $T(x)$ in terms of that for N:

$$\Pr[T(x) > t] = \Pr[N(x+t) - N(x) = 0] = e^{-\int_x^{x+t} \lambda(z)\,dz}.$$

If you've studied the survival model material in the Exam MLC/3 syllabus on life contingencies, then the integral above should look familiar: if you replace λ by μ, you get the formula for the survival probability ${}_tp_x$ with force of mortality μ. See, for example, Equation (3.2.14) in the *Actuarial Mathematics* textbook. Thus you can think of $\lambda(t)$ as the force with which nature is seeking to make an event occur at time t—which is why λ is also called the *intensity* function. This analogy with survival models can be helpful; for example, it allows you to view $\mathrm{E}[T(x)]$ as $\mathring{e}_x$ from life contingencies, which you know how to compute via $\mathring{e}_x = \int_0^\infty {}_tp_x\,dt$. That is,

$$\mathrm{E}[T(x)] = \int_0^\infty e^{-\int_x^{x+t} \lambda(z)\,dz}\,dt. \quad \P$$

Problems 1.1

[See my Foreword on page 2 for the web links.]

From the SoA Exam MLC/M/3 archives: Spring 2007 #5, 25, 26; Fall 2006 #10; Spring 2005 #25; Fall 2004 #16, 26; Fall 2003 #26; Fall 2002 #9; Fall 2001 #10.

From the CAS Exam 3 archives: Spring 2007 #1; Fall 2006 #28; Spring 2006 #33; Fall 2005 #26, 28; Spring 2005 #14; Fall 2004 #18, 19; Spring 2004 #27.

1.2 Compound Poisson processes

When employers provide health insurance to their employees, they are of course concerned about claim ***frequency***, the random number of claims filed. And they're concerned about claim ***severity***, the random size of each claim. But they are especially concerned about ***aggregate*** claims, the sum total of all the claims. This is the sum of a random number of random variables, and as such is extremely complicated to analyze; such a probability distribution is called a ***compound distribution***. If frequency is assumed to follow a Poisson process and the severities are independent and all have the same probability distribution, the result is a ***compound Poisson process***.

KEY ⇒ **Definition 1.8 (compound Poisson process)** *A compound Poisson process S has the following properties:*

1. *For $t > 0$, $S(t) = \sum_{j=1}^{N(t)} X_j$;*

2. *N is a Poisson process with rate function λ;*

3. *all the random variables X_j have the same distribution as a single random variable X;*

4. *for all t, the random variables $N(t)$ and all the X_j form an independent set; and*

5. *if $N(t)$ equals 0 for a particular value of t, then the empty sum $S(t)$ is taken to equal 0—so, in particular, $S(0) = 0$.*

Example 1.9 Suppose that health claims are filed with a health insurer at the Poisson rate $\lambda = 20$ per day, and that the independent severities X of each claim are Exponential random variables with mean $\theta = 500$. Then the aggregate S of claims is a compound Poisson process. ¶

Example 1.10 Suppose that automobile accident claims are filed with an automobile insurer at the Poisson rate $\lambda = 5$ per hour, and that the independent numbers X of people seriously injured in each accident are Binomial random variables with parameters 8 and 0.2. Then the aggregate number S of those seriously injured is a compound Poisson process. ¶

As I already indicated, compound Poisson processes are very complicated and difficult to analyze, and so approximation methods are often used. You should recall from basic probability that the sum of a "large" number of identically distributed independent random variables like our X_j is approximately a Normal random variable with the same mean and variance as the sum has. The same holds for the behavior of $S(t)$, provided that the expected value of $N(t)$ is "large". To approximate $S(t)$ by a Normal random variable, you need to be able to compute the mean and variance of $S(t)$.

From basic probability you know the iterated-expectation method for computing the mean of an expression $g(A, B)$ when A and B are themselves random:

$$\mathrm{E}[g(A,B)] = \mathrm{E}\big[\,\mathrm{E}[g(A,B) \mid B]\big],$$

where the inner expected value on the right of the equals sign is with respect to A, given B, and the outer expected value is with respect to B. The variance is more complicated:

$$\mathrm{Var}[g(A,B)] = \mathrm{E}\big[\,\mathrm{Var}[g(A,B) \mid B]\big] + \mathrm{Var}\big[\,\mathrm{E}[g(A,B) \mid B]\big],$$

where the inner expected value and variance are with respect to A, given B, and the outer expected value and variance are with respect to B. If you apply this approach to $S(t)$ with $N(t)$ playing the role of B and the various X_j playing the role of A, you get the following important formulas for the mean and variance of $S(t)$.

Fact 1.11 (compound-Poisson-process expectation and variance) *The following formulas hold for computing the expected value and variance of the compound-Poisson-process values $S(t)$ as defined in Key Definition 1.8 on compound Poisson processes:* ⇐ **KEY**

1. $\mathrm{E}[S(t)] = \mathrm{E}[N(t)]\,\mathrm{E}[X]$*; and*

2. $\mathrm{Var}[S(t)] = \mathrm{E}[N(t)]\,\mathrm{Var}[X] + \mathrm{Var}[N(t)]\,\mathrm{E}[X]^2 = \mathrm{E}[N(t)]\,\mathrm{E}[X^2]$.

In Key Fact 1.11 on compound-Poisson-process expectation and variance, the equation 1) should be intuitively clear—to get the average total claims, just multiply the average number of claims by the average size of each claim. The equation 2) in Key Fact 1.11 is less intuitive. The second equality in 2) follows from the facts that the mean and variance of the Poisson random variable $N(t)$ are equal, and that the variance of X plus the squared mean of X is simply the second moment $\mathrm{E}[X^2]$ of X.

Example 1.12 Consider the compound Poisson process in Example 1.9 modeling aggregate health claims; frequency N is a Poisson process with rate $\lambda = 20$ per day and severity X is an Exponential random variable with mean $\theta = 500$. Suppose that you are interested in the aggregate claims $S(10)$ during the first 10 days. You compute the mean of $S(10)$ using Key Fact 1.11 as follows:

$$\mathrm{E}[S(10)] = \mathrm{E}[N(10)]\,\mathrm{E}[X] = (20 \times 10)(500) = 100{,}000.$$

Where did the term 20×10 come from? Recall from Key Definition & Properties 1.2 on Poisson processes that $N(10) = N(10) - N(0)$ is a Poisson random variable M with mean $\Lambda = \int_0^{10} 20\,dz = 20 \times 10 = 200$. The variance of $N(10)$ is also 200, of course, so using the fact that $\mathrm{Var}[X] = \theta^2$ you compute the variance of $S(10)$ using Key Fact 1.11 as

$$\mathrm{Var}[S(10)] = \mathrm{E}[N(10)]\,\mathrm{Var}[X] + \mathrm{Var}[N(10)]\,\mathrm{E}[X]^2 = (200)(500^2) + (200)(500)^2 = 100{,}000{,}000.$$

Suppose now that you want to estimate the probability $\Pr[S(10) > 120{,}000]$ that aggregate claims in the first 10 days exceed 120,000. Approximate $S(10)$ by the Normal random variable $\mathcal{N}(100{,}000,\, 100{,}000{,}000)$ with the same mean 100,000 and variance $100{,}000{,}000 = 10{,}000^2$ as $S(10)$ has; here I use the notation $\mathcal{N}(\mu,\, \sigma^2)$ for the Normal random variable with mean μ and variance σ^2. The usual Normal-variable approximation method gives

$$\begin{aligned}
\Pr[S(10) > 120{,}000] &\approx \Pr[\mathcal{N}(100{,}000,\, 10{,}000^2) > 120{,}000] \\
&\approx \Pr\left[\frac{\mathcal{N}(100{,}000,\, 10{,}000^2) - 100{,}000}{10{,}000} > \frac{120{,}000 - 100{,}000}{10{,}000}\right] \\
&\approx \Pr[\mathcal{N}(0,1) > 2] \approx 0.0228
\end{aligned}$$

according to tables for the standard Normal random variable $\mathcal{N}(0,1)$. ¶

Problems 1.2

[See my Foreword on page 2 for the web links.]

From the SoA Exam MLC/M/3 archives: Fall 2005 #40; Fall 2002 #15; Spring 2001 #4.

From the CAS Exam 3 archives: Spring 2007 #2; Spring 2004 #26; Fall 2003 #30.

1.3 Creating new processes from existing ones

This Section examines three situations that arise in applications and result in creating new Poisson processes from existing ones—counting special types of events, adding two processes, and mixing processes. Although there are some connections among these situations, I'll present them one at a time.

1.3.1 Counting special types of events

In practice you may well be interested in counting various special types of the events being counted by a Poisson process, called "thinning" the Poisson process. For example, if you are counting insurance claims filed with an insurer as your basic events, you might be interested in counting those with claims of at most \$1,000, those greater than \$1,000 but at most \$10,000, and those greater than \$10,000. If claims are being filed at the Poisson rate of 100 per day and the probability of a claim's exceeding \$10,000 equals 0.2, so that 20% of claims on average exceed \$10,000, then it seems plausible that such large claims would be filed at a rate of $0.2 \times 100 = 20$ per day. What is not at all obvious, however, nor simple to prove, is that 20 is a ***Poisson*** rate—that is, that the number of these large claims constitutes a Poisson process with rate 20 per day. Since the proof of this result is beyond the scope of this note, I'll simply state the result.

KEY ⇒ Fact 1.13 (special types in Poisson processes, or thinning) *Suppose that the Poisson process* N *with rate function* λ *counts events, and that some of these events are viewed as special of various disjoint types (so no event can be of two different types): Type 1, Type 2, ..., Type* k*. For* $j = 1, 2, \ldots, k$ *let* $\pi_j(t)$ *denote the probability that an event occurring at time* t *is of Type* j*, and let* $\widetilde{N}_j$ *denote the counting process counting the events of Type* j*. Then* $\widetilde{N}_j$ *is a Poisson process with rate function* $\widetilde{\lambda}_j$ *given by* $\widetilde{\lambda}_j(t) = \pi_j(t)\lambda(t)$*. Moreover, for any set of positive numbers* $t_1, t_2, \ldots, t_k$*, the random variables* $\widetilde{N}_1(t_1), \widetilde{N}_2(t_2), \ldots, \widetilde{N}_k(t_k)$ *form an independent set.*

Example 1.14 Suppose that taxis depart an airport at the Poisson rate of 2 per minute and that the probabilities of a taxi having one, two, or three passengers are 0.5, 0.3, and 0.2, respectively. Suppose we are told that five taxis depart with one passenger in a 10-minute time interval from time t to time $t+10$; let's see what we can conclude about the random numbers of taxis departing with two or with three passengers in that time interval.

For $j = 1, 2, 3$, I'll call a departing taxi of Type j if it contains j passengers; in the notation of Key Fact 1.13 on special types in Poisson processes, we have $\pi_1 = 0.5$, $\pi_2 = 0.3$, and $\pi_3 = 0.2$. By Key Fact 1.13, the numbers $\widetilde{N}_1$, $\widetilde{N}_2$, and $\widetilde{N}_3$ of taxis departing with one, two, or three passengers are independent Poisson processes with rates $\widetilde{\lambda}_1 = 0.5 \times 2 = 1$, $\widetilde{\lambda}_2 = 0.6$, and $\widetilde{\lambda}_3 = 0.4$ per minute, respectively. I know that $\widetilde{N}_1(t+10) - \widetilde{N}_1(t) = 5$, and I want to know what we can conclude about $\widetilde{N}_2(t+10) - \widetilde{N}_2(t)$ and $\widetilde{N}_3(t+10) - \widetilde{N}_3(t)$.

First of all, since the $\widetilde{N}_j$ processes are ***independent***, the information about $\widetilde{N}_1$ in no way affects the behavior of the other two—they still behave as do all increments of Poisson processes. I'll just consider $\widetilde{N}_2$, the number of taxis departing with two passengers. By Key Definition & Properties 1.2 on Poisson processes, the increment $\widetilde{N}_2(t+10) - \widetilde{N}_2(t)$ is a Poisson random variable, say $\widetilde{M}_2$, with mean $\widetilde{\Lambda}_2 = \int_t^{t+10} \widetilde{\lambda}_2(z)\,dz = \int_t^{t+10} 0.6\,dz = 6$. [What about $\widetilde{N}_3(t+10) - \widetilde{N}_3(t)$?] So the mean of the number of taxis leaving with two passengers in that (or any other!) 10-minute period is 6, the variance is 6, and the probability that exactly three depart is $e^{-6}6^3/3! = 0.089235$. ¶

1.3.2 Sums of Poisson processes

You might sometimes want to add together two (or more) independent Poisson processes, where by independent I mean that any two increments, one from each process, are independent random variables. For example, N_1 might be counting the number of retirements from Company #1, which occur at the Poisson rate 10 per year, and N_2 might be counting those from Company #2, which occur at the Poisson rate 15 per year; for some reason you are interested in the total retirements from the two companies combined. It seems plausible that the combined retirements occur at the rate $10 + 15 = 25$ per year. What is not at all obvious, however, nor simple to prove, is that 25 is a ***Poisson*** rate—that is, that the combined number of retirements constitutes a Poisson process with rate 25 per year. Since the proof of this result is beyond the scope of this note, I'll simply state the result.

Fact 1.15 (sum of Poisson processes) *Suppose that the Poisson processes* N_1 *with rate function* λ_1 *and* N_2 *with rate function* λ_2 *are independent. Then the counting process* N *defined by* $N(t) = N_1(t) + N_2(t)$ *is a Poisson process with rate function* λ *given by* $\lambda(t) = \lambda_1(t) + \lambda_2(t)$*.*

1.3.3 Mixture distributions

New stochastic processes can be created as a ***mixture*** of Poisson processes. Since mixture distributions are important in other contexts as well and may not be familiar to you, please allow me to ramble a bit about mixtures in general. It might help to have a simple example first.

Example 1.16 Imagine a box full of fair dice; 20% of the dice are four-sided with the faces numbered one through four, and 80% are six-sided with the faces numbered one through six. If you repeatedly roll a four-sided die, you'll be seeing values from a Uniform random variable on the integers 1, 2, 3, 4, all equally likely. And if you repeatedly roll a six-sided die, you'll be seeing values from a Uniform random variable on the integers 1, 2, 3, 4, 5, 6, all equally likely.

But suppose that you reach into the box, grab a die at random (so a 20% chance of grabbing a four-sided die), roll it once, record the number, and return the die to the box. And then you repeat that process over and over, grabbing and rolling a random die each time. The results you will see are not from a Uniform random variable on 1, 2, 3, 4, nor from a Uniform random variable on 1, 2, 3, 4, 5, 6. Rather, they are from a 20/80 ***mixture*** of those two distributions.

Suppose that you want to compute the probability of rolling a 3 with a randomly chosen die—that is, $\Pr[N = 3]$ where N follows the mixture distribution. The rigorous mathematical computation of this proceeds as follows:

$$\begin{aligned}
\Pr[N = 3] &= \Pr[N = 3 \text{ and four-sided}] + \Pr[N = 3 \text{ and six-sided}] \\
&= \Pr[\text{four-sided}] \Pr[N = 3 \,|\, \text{four-sided}] + \Pr[\text{six-sided}] \Pr[N = 3 \,|\, \text{six-sided}] \\
&= (0.2)\left(\frac{1}{4}\right) + (0.8)\left(\frac{1}{6}\right) = \frac{11}{60} = 0.18333.
\end{aligned}$$

The more intuitive Mixing Method approach is as follows: pretend you know which case you are in and compute the answers [1/4 and 1/6 in this case], and then compute the expected value of those answers as the cases vary [so $(0.2)(1/4) + (0.8)(1/6) = 11/60$ as above]. Let's use this approach to compute the expected value of N. If you knew you had a four-sided die, the expected value would be 2.5, while it would be 3.5 for a six-sided die. So what is the expected value for a random die? Just the expected value of those two answers as the die-type varies: $(0.2)(2.5) + (0.8)(3.5) = 3.3$. ¶

For the Mixing Method used in Example 1.16 to be valid, the quantity you are computing for the mixture distribution must involve only a ***linear*** computation with the probability function (or density function for a continuous-type random variable) of the mixture—that is, no squaring of probabilities, no dividing by probabilities, et cetera. This is certainly true for an unconditional probability for the mixture, as just illustrated with $\Pr[N = 3]$ above. It's also valid for the expected value, which is just a sum of probabilities times values of the variable. Likewise for, say, the second moment, since that is just a sum of probabilities times squared values of the variable; in Example 1.16, for example, $\mathrm{E}[N^2] = (0.2)(30/4) + (0.8)(91/6) = 409/30 = 13.6333$. *But not so for the variance,* since the variance involves the square of the mean and thus the squaring of probabilities. In Example 1.16, for instance, $\mathrm{Var}[N]$ is correctly computed as the mixture's second moment 409/30 minus the square of the mixture's mean 3.3: $\mathrm{Var}[N] = 409/30 - 3.3^2 = 2.7433$. Mixing the two variances 15/12 and 35/12 in the 20/80 proportion gives the incorrect value 2.5833.

KEY ⇒ **Fact 1.17 (Mixing Method)** *Suppose you want to compute a quantity for a mixture distribution, and that quantity involves only linear computations with the probability function or density function of the mixture distribution. Then that quantity can be computed using the Mixing Method:*

1. *pretend that you know which case of the mixture holds, and compute the quantity; and*

2. *compute the expected value of the quantity as the case varies randomly.*

You already saw this method applied in Example 1.16 and its following paragraph. Here are some more examples.

Example 1.18 Given the value of Λ, $M \,|\, \Lambda$ is a Poisson random variable with mean Λ. But Λ is itself random, with $\Pr[\Lambda = 2] = 0.2$ and $\Pr[\Lambda = 3] = 0.8$. Let's find $\Pr[M = 4]$ and $\mathrm{E}[M]$ for the mixture distribution M using the Mixing Method. If we pretend we know the value of Λ, then

$\Pr[M = 4 \,|\, \Lambda] = e^{-\Lambda}\Lambda^4/4!$ and $\mathrm{E}[M \,|\, \Lambda] = \Lambda$. We get the values for the mixture distribution by taking the expectations of these values as Λ varies:

$$\begin{aligned}\Pr[M = 4] &= (0.2)\left(e^{-2}2^4/4!\right) + (0.8)\left(e^{-3}3^4/4!\right) = 0.15247, \text{ and}\\ \mathrm{E}[M] &= (0.2)(2) + (0.8)(3) = 2.8. \quad \P\end{aligned}$$

Example 1.19 Given the value of Λ, $M \,|\, \Lambda$ is a Poisson random variable with mean Λ. But Λ is itself random, namely a Uniform random variable on $[2, 3]$. Let's find $\Pr[M = 4]$ and $\mathrm{E}[M]$ for the mixture distribution M using the Mixing Method. If we pretend we know the value of Λ, then $\Pr[M = 4 \,|\, \Lambda] = e^{-\Lambda}\Lambda^4/4!$ and $\mathrm{E}[M \,|\, \Lambda] = \Lambda$. We get the values for the mixture distribution by taking the expectations of these values as Λ varies:

$$\begin{aligned}\Pr[M = 4] &= \mathrm{E}[e^{-\Lambda}\Lambda^4/4!] = \int_2^3 f_\Lambda(z)e^{-z}z^4/4!\,dz = \int_2^3 \frac{1}{3-2}e^{-z}z^4/4!\,dz\\ &= 0.13208,\end{aligned}$$

where the final numerical result follows from a tedious handful of integrations by parts. The mean is much simpler:

$$\mathrm{E}[M] = \mathrm{E}[\Lambda] = \frac{2+3}{2} = 2.5. \quad \P$$

Since the next example makes use of both the Gamma distribution and the Negative Binomial distribution, I want to remind you about those first.

Definition & Properties 1.20 (Gamma random variable) *A Gamma random variable X with two parameters $\alpha > 0$ and $\theta > 0$ is a continuous-type random variable $X > 0$ with*

1. *its density function f_X given by $f_X(x) = cx^{\alpha-1}e^{-\frac{1}{\theta}x}$ for some positive constant c.*

   From this it follows that

2. $\mathrm{E}[X] = \alpha\theta$*;*
3. $\mathrm{Var}[X] = \alpha\theta^2$*;*
4. *if the parameter α is a positive integer, then X has the same distribution as the sum of α independent Exponential random variables, each having mean θ;*
5. *when $\alpha = 1$, X is an Exponential random variable with mean θ; and*
6. *for any positive constant C, the random variable $\widetilde{X} = CX$ is a Gamma random variable with parameters $\widetilde{\alpha} = \alpha$ and $\widetilde{\theta} = C\theta$.*

Definition & Properties 1.21 (Negative Binomial random variable) *A Negative Binomial random variable M with parameters $r > 0$ and $\beta > 0$ is a discrete-type random variable for which:*

1. *its only possible values are the non-negative integers 0, 1, 2, 3, ...; and*
2. $\Pr[M = k] = \frac{r(r+1)\cdots(r+k-1)}{k!}\frac{\beta^k}{(1+\beta)^{k+r}}$.

   From these it follows that

3. $\mathrm{E}[M] = r\beta$*; and*
4. $\mathrm{Var}[M] = r\beta(1+\beta)$.

Example 1.22 Given the value of Λ, $M \mid \Lambda$ is a Poisson random variable with mean Λ. But Λ is itself random, namely a Gamma random variable with parameters $\alpha = 1.5$ and $\theta = 2$. Let's again try to find $\Pr[M = 4]$ and $\mathrm{E}[M]$ for the mixture distribution M using the Mixing Method. If we pretend we know the value of Λ, then $\Pr[M = 4 \mid \Lambda] = e^{-\Lambda}\Lambda^4/4!$ and $\mathrm{E}[M \mid \Lambda] = \Lambda$. The Mixing Method for the mean goes quite simply: $\mathrm{E}[M] = \mathrm{E}[\Lambda] = \alpha\theta = 1.5 \times 2 = 3$. The Mixing Method for the probability presents the unpleasant task of computing the expected value of $e^{-\Lambda}\Lambda^4/4!$ for a Gamma random variable Λ. Fortunately—and here's why I reminded you about the Negative Binomial distribution—it is known (and you'll see the details of this when you study for SoA/CAS Exam C/4) that

> the mixture distribution in this general setting (Poisson and Gamma distributions) is always a Negative Binomial random variable with parameters given by $r = \alpha$ and $\beta = \theta$.

So, in our case that's $r = 1.5$ and $\beta = 2$. Note by property 3) in Definition & Properties 1.21 on Negative Binomial random variables that this again gives $\mathrm{E}[M] = r\beta = 1.5 \times 2 = 3$, just as by the Mixing Method. But now we can also easily compute $\Pr[M = 4] = \frac{1.5\times2.5\times3.5\times4.5}{4!}\frac{2^4}{(1+2)^{4+1.5}} = 0.093552$. ¶

You may well be wondering why I've rambled on for so long about mixture distributions for Poisson random variables when this note is primarily about Poisson *processes*. My answer is that you also encounter mixtures of Poisson *processes*, and the key to handling them is the ability to handle mixtures of Poisson random variables. How might mixtures of Poisson processes arise? An insurer, for example, might model the number of claims filed by each individual insured as a homogeneous Poisson process with rate λ, but believe that λ should be viewed as varying randomly among insureds—and that's a mixture.

Example 1.23 Suppose that, given the value of λ, N is a homogeneous Poisson process with rate λ, but that λ itself is random, namely a Uniform random variable on $[1, 1.5]$. Let's find $\Pr[N(2) = 4]$ and $\mathrm{E}[N(2)]$ for the mixed process—that is, for random values of $N(2)$ from random λ's. If we pretend that we know λ, then N is a Poisson process with rate λ and so $N(2)$—which equals the increment $N(2) - N(0)$— is a Poisson random variable M with mean $\Lambda = \int_0^2 \lambda\,dz = 2\lambda$. Since λ's values are uniformly distributed over $[1, 1.5]$, it's clear that Λ's values, which are double λ's, are uniformly distributed over $[2, 3]$. That is, exactly as in Example 1.19, $M \mid \Lambda$ is a Poisson random variable with mean Λ while Λ is a Uniform random variable on $[2, 3]$. So, as in Example 1.19, $\Pr[N(2) = 4] = \Pr[M = 4] = 0.13208$. Likewise, $\mathrm{E}[N(2)] = \mathrm{E}[M] = 2.5$ as in Example 1.19. ¶

Example 1.24 Suppose that, given the value of λ, N is a homogeneous Poisson process with rate λ, but that λ itself is random, namely a Gamma random variable with parameters $\alpha = 1.5$ and $\theta = 1$. Let's find $\Pr[N(2) = 4]$ and $\mathrm{E}[N(2)]$ for the mixed process—that is, for random values of $N(2)$ from random λ's. If we pretend that we know λ, then N is a Poisson process with rate λ and so $N(2)$—which equals the increment $N(2) - N(0)$— is a Poisson random variable M with mean $\Lambda = \int_0^2 \lambda\,dz = 2\lambda$. Since λ is a Gamma random variable with parameters $\alpha = 1.5$ and $\theta = 1$, property 6) of Definition & Properties 1.20 on Gamma random variables tells us that 2λ—which is Λ—is a Gamma random variable with parameters $\widetilde{\alpha} = \alpha = 1.5$ and $\widetilde{\theta} = 2\theta = 2$. That is, exactly as in Example 1.22, $M \mid \Lambda$ is a Poisson random variable with mean Λ while Λ is a Gamma random variable with parameters 1.5 and 2. So, as in Example 1.22, $\Pr[N(2) = 4] = \Pr[M = 4] = 0.093552$. Likewise, $\mathrm{E}[N(2)] = \mathrm{E}[M] = 3$ as in Example 1.22. ¶

Problems 1.3

[See my Foreword on page 2 for the web links.]

From the SoA Exam MLC/M/3 archives: Spring 2007 #6; Fall 2006 #9; Spring 2005 #5, 6; Fall 2002 #5, 20; Fall 2001 #19, 27; Spring 2001 #3, 15, 37.

From the CAS Exam 3 archives: Fall 2005 #27, 31; Spring 2005 #7, 17; Fall 2004 #17; Spring 2004 #31; Fall 2003 #31.

1.4 Homogeneous Poisson processes

This section focuses on *homogeneous* Poisson processes: Poisson processes for which the rate function λ is a constant. Everything in the previous three sections about general Poisson processes remains true of course for homogeneous ones, but some things are simpler in the homogeneous case.

1.4.1 Fundamentals

Consider first the statement about the distribution of increments in property 3) of Key Definition & Properties 1.2 on Poisson processes. In the homogeneous case, as you have probably realized yourself by now from Examples 1.12, 1.14, 1.23, and 1.24, the mean Λ of the increment $N(t+h) - N(t)$ is $\Lambda = \int_t^{t+h} \lambda \, dz = \lambda h$, simply the constant rate λ times the length of the time interval over which the increment is defined. Note that Λ therefore does not depend on *where* the interval of length h lies, but only on its *length*; this is described by saying that homogeneous Poisson processes have ***stationary increments***. This simple formula for Λ deserves highlighting:

Fact 1.25 (increments in a homogeneous Poisson process) *If N is a homogeneous Poisson process with rate λ, then each increment $N(t+h) - N(t)$ is a Poisson random variable with mean $\Lambda = \lambda h$.* $\Leftarrow$ **KEY**

The behavior of the important time intervals T_n and V_j in Definition 1.4 on time intervals is simpler in the homogeneous case as well. Consider the first interevent time V_1, for example; this is the same as the first event time T_1. I'll use the approach demonstrated in Example 1.6 to show that V_1 is an Exponential random variable with mean $\theta = \frac{1}{\lambda}$. [Note that having the expected interevent time equal $\frac{1}{\lambda}$ is plausible—if events occur on average at the rate $\lambda = 5$ per year, say, then it seems plausible that they are on averge $\frac{1}{5}$ of a year apart.]

The argument for V_1 is computational: $\Pr[V_1 > t] = \Pr[N(t) = 0] = e^{-\Lambda} = e^{-\lambda t}$. This says that the cumulative distribution function F_{V_1} for V_1 is given by $F_{V_1}(t) = 1 - e^{-\lambda t}$, which is the cumulative distribution function for an Exponential random variable with mean $\theta = \frac{1}{\lambda}$ as asserted. So that means that V_1 *is* an Exponential random variable with mean $\theta = \frac{1}{\lambda}$.

A slightly more complicated argument shows that not just V_1 but all of the V_j are Exponential random variables with mean $\theta = \frac{1}{\lambda}$, and moreover that they together form an independent set. Since the n^{th} event time T_n satisfies $T_n = V_1 + V_2 + \cdots + V_n$, property 4) of Definition & Properties 1.20 on Gamma Random variables shows that T_n is a Gamma random variable with parameters $\alpha = n$ and $\theta = \frac{1}{\lambda}$. These facts deserve emphasis.

Fact 1.26 (V_j and T_n for homogeneous Poisson processes) *Suppose that N is a homogeneous Poisson process with constant rate λ. Then* $\Leftarrow$ **KEY**

1. *the interevent times V_j are independent Exponential random variables, each with mean $\theta = \frac{1}{\lambda}$; and*

2. *the event times T_n are Gamma random variables with parameters $\alpha = n$ and $\theta = \frac{1}{\lambda}$.*

The fact that interevent times are Exponential random variables with mean $\frac{1}{\lambda}$ provides a way to think about homogeneous Poisson processes. Remember that Example 1.7 showed that the intensity function λ can be viewed as the force with which nature is trying to make events occur, just as in the survival model material in life contingencies μ is the force of mortality. I imagine that your favorite model in that survival models material is the forever-constant-force (or Exponential) model, in which the force of mortality μ is a constant. Now you can see that a homogeneous Poisson process is just counting events subject to a constant force λ.

Surprisingly, property 1) of Key Fact 1.26 on V_j and T_n for homogeneous processes is ***equivalent*** to N's being a homogeneous Poisson process, although the proof is beyond the scope of this note.

KEY ⇒ **Fact 1.27 (Exponential interevent times and homogeneous processes)** *Suppose that N is a counting process—$N(0) = 0$, and for $t > 0$ $N(t)$ equals the random number of events of interest occurring after time 0 and by time t. Suppose also that the interevent times V_j between successive events are independent Exponential random variables each having mean θ. Then N is a homogeneous Poisson process with constant rate $\lambda = \frac{1}{\theta}$.*

1.4.2 Sums of compound Poisson processes

Suppose that S_1 models aggregate claims from Insurer #1 with severity X_1 and is a compound Poisson process as in Key Definition 1.8 on compound Poisson processes, with

$$S_1(t) = \sum_{j=1}^{N_1(t)} X_{1,j}.$$

Suppose also that S_2 models aggregate claims for Insurer #2 with severity X_2 and is a compound Poisson process, with

$$S_2(t) = \sum_{j=1}^{N_2(t)} X_{2,j}.$$

Suppose further that, for each $t > 0$, $N_1(t)$, $N_2(t)$, and all the $X_{1,j}$ and $X_{2,k}$ form an independent set. Then define S by $S(t) = S_1(t) + S_2(t)$; S models aggregate claims for a merged insurer.

You know from Fact 1.15 on sums of Poisson processes that the total number N of claims as given by $N(t) = N_1(t) + N_2(t)$ is itself a Poisson process with rate $\lambda = \lambda_1 + \lambda_2$, so it seems that S itself is a compound Poisson process. But be careful—what's the severity X in S? Sometimes the severity is X_1 and sometimes it's X_2, which sounds like a mixture distribution.

I'll argue intuitively. Severities X_1 occur on average at the rate λ_1, severities X_2 occur on average at the rate λ_2, and the mixed severities X occur on average at the rate $\lambda = \lambda_1 + \lambda_2$. So, on average, λ_1 out of the $\lambda_1 + \lambda_2$ severities per time period come from X_1, while λ_2 out of the $\lambda_1 + \lambda_2$ severities per time period come from X_2. It thus seems plausible (and in fact can be proved) that X is in fact a mixture of the two distributions X_1 and X_2 with probabilities $\frac{\lambda_1}{\lambda_1+\lambda_2}$ and $\frac{\lambda_2}{\lambda_1+\lambda_2}$, respectively. Stated differently, $\frac{\lambda_1}{\lambda_1+\lambda_2}$ is the probability, at any point in time, that the next event counted by the summed Poisson process N will in fact come from the Poisson process N_1.

Example 1.28 In the notation of this subsection, suppose that $\lambda_1 = 5$ and $\lambda_2 = 15$. Suppose also that X_1 is an Exponential random variable with mean $\theta_1 = 50$ and X_2 is an Exponential random variable with mean $\theta_2 = 20$. Then N is a Poisson process with rate $\lambda = 5 + 15 = 20$ and X is a mixture of X_1 and X_2 with probabilities $\frac{5}{5+15} = 0.25$ and 0.75, respectively. ¶

1.4.3 Counting special types of events in a homogeneous process

Go back and review Key Fact 1.13 on special types in Poisson processes, or thinning. I'll wait while you do so .... Now suppose that the original Poisson process N is *homogeneous* so that λ is a constant. Then whether the rate function $\widetilde{\lambda}_j$ of the Poisson process $\widetilde{N}_j$ counting events of Type j is a constant depends on whether the probability $\pi_j(t)$ of an event at time t being special of Type j is a constant. That is, the Poisson process $\widetilde{N}_j$ could be either homogeneous or non-homogeneous.

In practice, one of the ways that non-homogeneous Poisson processes arise is by counting special events from a homogeneous Poisson process.

Example 1.29 Suppose that the number of insurance claims filed with an insurer is modeled by a homogeneous Poisson process N with rate $\lambda = 50$ per hour, and the insurer is interested in the "special" claims for which the cost of repair exceeds \$1,000. Because of inflation in the cost of repairs, the probability $\pi(t)$ of the cost's exceeding \$1,000 for a claim at time t probably increases with time; suppose that $\pi(t) = 0.8 - 0.1e^{-t}$. Then the number of these special claims is modeled by the non-homogeneous Poisson process $\widetilde{N}$ with rate function $\widetilde{\lambda}$ given by $\widetilde{\lambda}(t) = \pi(t)\lambda = 50(0.8 - 0.1e^{-t})$.

Problems 1.4

[See my Foreword on page 2 for the web links.]

From the SoA Exam MLC/M/3 archives: Spring 2007 #5, 6, 26; Fall 2006 #8, 9; Fall 2005 #7,8; Spring 2005 #5, 6, 24; Fall 2003 #11, 20, 26; Fall 2002 #9.

From the CAS Exam 3 archives: Spring 2007 #1, 2; Fall 2006 #26, 27; Fall 2005 #29, 31; Spring 2005 #7, 11, 17; Fall 2004 #17; Spring 2004 #31; Fall 2003 #13, #31.