

14.5: Thinning and Superposition

Thinning

Thinning or *splitting* a Poisson process refers to classifying each random point, independently, into one of a finite number of different types. The random points of a given type also form Poisson processes, and these processes are independent. Our exposition will concentrate on the case of just two types, but this case has all of the essential ideas.

The Two-Type Process

We start with a Poisson process with rate $r > 0$. Recall that this statement really means three interrelated stochastic processes: the sequence of inter-arrival times $\mathbf{X} = (X_1, X_2, \dots)$, the sequence of arrival times $\mathbf{T} = (T_0, T_1, \dots)$, and the counting process $\mathbf{N} = (N_t : t \geq 0)$. Suppose now that each arrival, independently of the others, is one of two types: type 1 with probability p and type 0 with probability $1 - p$, where $p \in (0, 1)$ is a parameter. Here are some common examples:

- The arrivals are radioactive emissions and each emitted particle is either *detected* (type 1) or *missed* (type 0) by a counter.
- The arrivals are customers at a service station and each customer is classified as either *male* (type 1) or *female* (type 0).

We want to consider the type 1 and type 0 random points separately. For this reason, the new random process is usually referred to as *thinning* or *splitting* the original Poisson process. In some applications, the type 1 points are *accepted* while the type 0 points are *rejected*. The main result of this section is that the type 1 and type 0 points form separate Poisson processes, with rates rp and $r(1 - p)$ respectively, and are independent. We will explore this important result from several points of view.

Bernoulli Trials

In the previous sections, we have explored the analogy between the Bernoulli trials process and the Poisson process. Both have the strong renewal property that at each fixed time and at each arrival time, the process stochastically “restarts”, independently of the past. The difference, of course, is that time is discrete in the Bernoulli trials process and continuous in the Poisson process. In this section, we have both processes simultaneously, and given our previous explorations, it's perhaps not surprising that this leads to some interesting mathematics.

Thus, in addition to the processes \mathbf{X} , \mathbf{T} , and \mathbf{N} , we have a sequence of Bernoulli trials $\mathbf{I} = (I_1, I_2, \dots)$ with success parameter p . Indicator variable I_j specifies the type of the j th arrival. Moreover, because of our assumptions, \mathbf{I} is independent of \mathbf{X} , \mathbf{T} , and \mathbf{N} . Recall that V_k , the trial number of the k th success has the negative binomial distribution with parameters k and p for $k \in \mathbb{N}_+$. We take $V_0 = 0$ by convention. Also, U_k , the number of trials needed to go from the $(k - 1)$ st success to the k th success has the geometric distribution with success parameter p for $k \in \mathbb{N}_+$. Moreover, $\mathbf{U} = (U_1, U_2, \dots)$ is independent and $\mathbf{V} = (V_0, V_1, \dots)$ is the partial sum process associated with \mathbf{U} :

$$V_k = \sum_{i=1}^k U_i, \quad k \in \mathbb{N} \quad (14.5.1)$$

$$U_k = V_k - V_{k-1}, \quad k \in \mathbb{N}_+ \quad (14.5.2)$$

As noted above, the Bernoulli trials process can be thought of as random points in *discrete* time, namely the trial numbers of the successes. With this understanding, \mathbf{U} is the sequence of inter-arrival times and \mathbf{V} is the sequence of arrival times.

The Inter-arrival Times

Now consider just the type 1 points in our Poisson process. The time between the arrivals of $(k - 1)$ st and k th type 1 point is

$$Y_k = \sum_{i=V_{k-1}+1}^{V_k} X_i, \quad k \in \mathbb{N}_+ \quad (14.5.3)$$

Note that Y_k has U_k terms. The next result shows that the type 1 points form a Poisson process with rate pr .

$\mathbf{Y} = (Y_1, Y_2, \dots)$ is a sequence of independent variables and each has the exponential distribution with rate parameter pr

Proof

From the renewal properties of the Poisson process and the Bernoulli trials process, the inter-arrival times are independent and identically distributed. Each inter-arrival time is the sum of a random number of independent terms; each term has the exponential distribution with rate r , and the number of terms has the geometric distribution on \mathbb{N}_+ with parameter p . Moreover, the number of terms is independent of the terms themselves. We showed in the section on the exponential distribution that a random sum of this form has the exponential distribution with parameter rp .

Similarly, if $\mathbf{Z} = (Z_1, Z_2, \dots)$ is the sequence of interarrival times for the type 0 points, then \mathbf{Z} is a sequence of independent variables, and each has the exponential distribution with rate $(1-p)r$. Moreover, \mathbf{Y} and \mathbf{Z} are independent.

Counting Processes

For $t \geq 0$, let M_t denote the number of type 1 arrivals in $(0, t]$ and W_t the number of type 0 arrivals in $(0, t]$. Thus, $\mathbf{M} = (M_t : t \geq 0)$ and $\mathbf{W} = (W_t : t \geq 0)$ are the counting processes for the type 1 arrivals and for the type 0 arrivals. The next result follows from the previous subsection, but a direct proof is interesting.

For $t \geq 0$, M_t has the Poisson distribution with parameter prt , W_t has the Poisson distribution with parameter $(1-p)r$, and M_t and W_t are independent.

Proof

The important observation is that the conditional distribution of M_t given $N_t = n$ is binomial with parameters n and p . Thus for $j \in \mathbb{N}$ and $k \in \mathbb{N}$,

$$\mathbb{P}(M_t = j, W_t = k) = \mathbb{P}(M_t = j, N_t = j+k) = \mathbb{P}(N_t = j+k) \mathbb{P}(M_t = j \mid N_t = j+k) \quad (14.5.4)$$

$$= e^{-rt} \frac{(rt)^{j+k}}{(j+k)!} \frac{(j+k)!}{j!k!} p^j (1-p)^k \quad (14.5.5)$$

$$= e^{-prt} \frac{(prt)^j}{j!} e^{-(1-p)rt} \frac{[(1-p)rt]^k}{k!} \quad (14.5.6)$$

In the two-type Poisson experiment vary r , p , and t with the scroll bars and note the shape of the probability density functions. For various values of the parameters, run the experiment 1000 times and compare the relative frequency functions to the probability density functions.

Estimating the Number of Arrivals

Suppose that the type 1 arrivals are observable, but not the type 0 arrivals. This setting is natural, for example, if the arrivals are radioactive emissions, and the type 1 arrivals are emissions that are detected by a counter, while the type 0 arrivals are emissions that are missed. Suppose that for a given $t > 0$, we would like to estimate the total number arrivals N_t after observing the number of type 1 arrivals M_t .

The conditional distribution of N_t given $M_t = k$ is the same as the distribution of $k + W_t$.

$$\mathbb{P}(N_t = n \mid M_t = k) = e^{-(1-p)rt} \frac{[(1-p)rt]^{n-k}}{(n-k)!}, \quad n \in \{k, k+1, \dots\} \quad (14.5.7)$$

Proof

Recall from the basic [splitting result above](#) that M_t and W_t are independent. Thus, for $n \in \mathbb{N}$,

$$\mathbb{P}(N_t = n \mid M_t = k) = \frac{\mathbb{P}(N_t = n, M_t = k)}{\mathbb{P}(M_t = k)} = \frac{\mathbb{P}(M_t = k, W_t = n-k)}{\mathbb{P}(M_t = k)} \quad (14.5.8)$$

$$= \frac{\mathbb{P}(M_t = k) \mathbb{P}(W_t = n-k)}{\mathbb{P}(M_t = k)} = \mathbb{P}(W_t = n-k) \quad (14.5.9)$$

The form of the probability density function follows since W_t as the Poisson distribution with parameter $(1-p)r$.

$$\mathbb{E}(N_t \mid M_t = k) = k + (1-p)r \quad .$$

Proof

This follows easily from our previous theorem since $\mathbb{E}(N_t | M_t = k) = \mathbb{E}(k + W_t) = k + (1 - p)r$.

Thus, if the overall rate r of the process and the probability p that an arrival is type 1 are known, then it follows from the general theory of conditional expectation that the best estimator of N_t based on M_t , in the least squares sense, is

$$\mathbb{E}(N_t | M_t) = M_t + (1 - p)r \quad (14.5.10)$$

The mean square error is $\mathbb{E}([N_t - \mathbb{E}(N_t | M_t)]^2) = (1 - p)rt$.

Proof

Note that $N_t - \mathbb{E}(N_t | M_t) = W_t - (1 - p)r$. Thus the mean square error is just $\text{var}(W_t) = (1 - p)rt$.

The Multi-Type Process

As you might guess, the results above generalize from 2 types to k types for general $k \in \mathbb{N}_+$. Once again, we start with a Poisson process with rate $r > 0$. Suppose that each arrival, independently of the others, is type i with probability p_i for $i \in \{0, 1, \dots, k-1\}$. Of course we must have $p_i \geq 0$ for each i and $\sum_{i=0}^{k-1} p_i = 1$. Then for each i , the type i points form a Poisson process with rate $p_i r$, and these processes are independent.

Superposition

Complementary to splitting or thinning a Poisson process is *superposition*: if we combine the random points in time from independent Poisson processes, then we have a new Poisson processes. The rate of the new process is the sum of the rates of the processes that were combined. Once again, our exposition will concentrate on the superposition of two processes. This case contains all of the essential ideas.

Two Processes

Suppose that we have two independent Poisson processes. We will denote the sequence of inter-arrival times, the sequence of arrival times, and the counting variables for the process $i \in \{1, 2\}$ by $\mathbf{X}^i = (X_1^i, X_2^i, \dots)$, $\mathbf{T}^i = (T_1^i, T_2^i, \dots)$, and $\mathbf{N}^i = (N_t^i : t \in [0, \infty))$, and we assume that process i has rate $r_i \in (0, \infty)$. The new process that we want to consider is obtained by simply combining the random points. That is, the new random points are $\{T_n^1 : n \in \mathbb{N}_+\} \cup \{T_n^2 : n \in \mathbb{N}_+\}$, but of course then ordered in time. We will denote the sequence of inter-arrival times, the sequence of arrival times, and the counting variables for the new process by $\mathbf{X} = (X_1, X_2, \dots)$, $\mathbf{T} = (T_1, T_2, \dots)$, and $\mathbf{N} = (N_t : t \in [0, \infty))$. Clearly if A is an interval in $[0, \infty)$ then

$$N(A) = N^1(A) + N^2(A) \quad (14.5.11)$$

the number of combined points in A is simply the sum of the number of point in A for processes 1 and 2. It's also worth noting that

$$X_1 = \min\{X_1^1, X_1^2\} \quad (14.5.12)$$

the first arrival for the combined process is the smaller of the first arrival times for processes 1 and 2. The other inter-arrival times, and hence also the arrival times, for the combined process are harder to state.

The combined process is a Poisson process with rate $r_1 + r_2$. Moreover,

Proof

As noted above, if A is a subinterval of $[0, \infty)$ then $N(A) = N^1(A) + N^2(A)$. The first term has the Poisson distribution with parameter $r_1 \lambda(A)$, the second term has the Poisson distribution with parameter $r_2 \lambda(A)$, and the terms are independent. Hence $N(A)$ has the Poisson distribution with parameter $r_1 \lambda(A) + r_2 \lambda(A) = (r_1 + r_2) \lambda(A)$. Thus the counting process has stationary, Poisson distributed increments. Next, if (A_1, A_2, \dots, A_n) is a sequence of disjoint subintervals of $[0, \infty)$ then

$$(N(A_1), N(A_2), \dots, N(A_n)) = (N^1(A_1) + N^2(A_1), N^1(A_2) + N^2(A_2), \dots, N^1(A_n) + N^2(A_n)) \quad (14.5.13)$$

is an independent sequence, so the counting process has independent increments.

Computational Exercises

In the two-type Poisson experiment, set $r = 2$, $t = 3$, and $p = 0.7$. Run the experiment 1000 times, Compute the appropriate relative frequency functions and investigate empirically the independence of the number of type 1 points and the number of type 0 points.

Suppose that customers arrive at a service station according to the Poisson model, with rate $r = 20$ per hour. Moreover, each customer, independently, is female with probability 0.6 and male with probability 0.4. Find the probability that in a 2 hour period, there will be at least 20 women and at least 15 men.

Answer

0.5814

In the two-type Poisson experiment, set $r = 3$, $t = 4$, and $p = 0.8$. Run the experiment 100 times.

1. Compute the estimate of N_t based on M_t for each run.
2. Over the 100 runs, compute average of the sum of the squares of the errors.
3. Compare the result in (b) with the result in Exercise 8.

Suppose that a piece of radioactive material emits particles according to the Poisson model at a rate of $r = 100$ per second. Moreover, assume that a counter detects each emitted particle, independently, with probability 0.9. Suppose that the number of detected particles in a 5 second period is 465.

1. Estimate the number of particles emitted.
2. Compute the mean square error of the estimate.

Answer

1. 515
2. 50

This page titled [14.5: Thinning and Superposition](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist](#) ([Random Services](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.