

We often think of a Bernoulli process as a stream of arriving traffic. What happens if we merge two streams? For example, consider a server that receives traffic from two independent sources. How do we describe the total traffic that arrives to this server?

Here's a precise model. We have two streams that correspond to Bernoulli processes with some parameters each, p and q , respectively. And each one of these processes receives arrivals at certain times that we indicate by crosses; and similarly, for the second process.

We assume that these two processes are independent. And what we mean by this is that any collection of random variables associated with the first process will be independent from any collection of random variables associated with the second process. We now merge the two processes as follows. Whenever there's an arrival in any of the original processes, we record an arrival in the merged process, as in this picture.

Notice that we do not make a distinction between those slots at which there was an arrival in only one of the original [processes] versus those slots in which there was an arrival in both of the original processes. In both cases, we just say that there was an arrival in the merged process, and so collisions-- arrivals in both of the original processes-- are counted as only one arrival in the merged process.

Now, what can we say about the merged process?

We will argue that it is a Bernoulli process with a certain parameter that we will compute. To check the Bernoulli property for the merged process, the first thing we need to ensure is the independence assumption, independence across slots. Let us look at two typical slots. And to do this, it helps to define some notation that X_t and Y_t be the original processes, and let Z_t be the merged process.

The random variable Z_t is determined in some way by the random variables X_t and Y_t . If I tell you there was an arrival in the first and to the second process, you can tell whether there was an arrival in the merged process. And similarly, the random variable Z_{t+1} is determined in some way from X_{t+1} and Y_{t+1} . What this is saying is that whether we have an arrival at this slot is determined by what happens in these two slots. And whether we have an arrival in this slot is determined by whatever happens in these two slots.

Now, we have assumed that the two processes are independent. So these two random variables are independent from those two random variables. And furthermore, across time, this random variable will be independent from that random variable. And this random variable will be independent from that random variable.

So these four random variables are independent of each other. Because of this, we have Z_t , a function of two random variables that are independent from the two random variables that determine Z_t plus 1. And for this reason, Z_t and Z_t plus 1 will be independent. This proves a pairwise independence property for the merged process, but we can extend this argument to argue that the collection of random variables, Z_1 up to Z_t , is a collection of independent random variables. So we have the independence property.

Now, let us calculate the probability of an arrival during a typical slot. During a typical time slot, there are four possibilities for what may occur. And these possibilities have to do with whether in the X process, we have an arrival or not; and in the Y process, whether we have an arrival or not.

The probability that we have an arrival in both processes, because of independence, is the product of the probability that we have an arrival in the first process with the probability that we have an arrival in the second process. Similarly, there's a probability p of an arrival in the first process and no arrival in the second. There's a probability $1 - p$ of no arrival in the first process and [a probability q of] an arrival in the second process. And finally, there is probability $1 - p$ times $1 - q$ of no arrival in either of the two processes.

The probability that we have an arrival in the merged process is the probability of this green event. These are the cases in which an arrival gets recorded in the merged process. So the probability of an arrival in the merged process is the sum of those three probabilities. Or another way to calculate it is 1 minus this probability here; namely, $1 - (1 - p)(1 - q)$.

And after you expand this product and do some cancellations, you end up with this expression, which is the probability of an arrival during a slot in the merged process. Of course, this probability is constant across time. And this, together with the independence property, establishes that the merged process is actually a Bernoulli process.

Now, let us end by answering one more question. If I tell you that at a certain time slot, there was at

least one arrival in the two processes, which means that there was an arrival in the merged process, what is the probability that there was an arrival in the first process? Now, the event that there was an arrival in the first process is this event here. So we're trying to calculate the conditional probability of the blue event given that the green event has occurred.

We use the definition of conditional probabilities. A conditional probability is equal to the probability that both events happen, which in this case is the intersection of the blue and the green event, which coincides with the blue event. And the probability of the blue event is the sum of these two numbers, is equal to p . And then we divide by the probability of the conditioning event, this is the probability of an arrival. But this is what we have just calculated, which is $p + q - p \times q$.