In this section, we only focus on how to generate RFID data using Poisson distribution. The reason of why Poisson distribution has been used is because it is commonly used to model the number of expected events for a process given we know the average rate at which events occur during a given unit of time. For example, assume that there is 100 RFID readings have been captured per second using one RFID reader. Can we generate a simulation of number of RFID readings for the next 10 seconds?

According to [ ], we can also describe this process as:

* A window of observation – a specific time period in which events can occur.
* A rate of occurrence – how often is an event expected to occur in that window
* The number of times an event occurs (the observation).

Theoretically, the Poisson distribution is the discrete probability distribution of the number of events occurring in a given time period, given the average number of times the event occurs over that time period [ ].

The Poisson distribution is applicable only when several conditions hold [ ].

* An event can occur any number of times during a time period.
* Events occur independently. In other words, if an event occurs, it does not affect the probability of another event occurring in the same time period.
* The rate of occurrence is constant; that is; the rate does not change based on time.
* The probability of an event occurring is proportional to the length of the time period. For example, it should be twice as likely for an event to occur in a 2 hour time period than it is for an event to occur in 1-hour period.

For example, the Poisson distribution is appropriate for modeling the number of phone calls an office would receive during the noon hour, if they know that they average 4 calls per hour during that time period.

* Although the average is 4 calls, they could theoretically get any number of calls during that time period.
* The events are effectively independently since they is no reason to expect a caller to affect the chances of another person calling.
* The occurrence rate may be assumed to be constant.
* It is reasonable to assume that (for example) the probability of getting a call in the first half hour is the same as the probability of getting a call in the final half hour.

Of course, this situation isn’t an absolute perfect theoretical fit for the Poisson distribution. For instance, the office certainly cannot receive a trillion calls during the time period, as there are less than a trillion people alive to be making phone calls. Practically speaking, the situation is close enough that the Poisson distribution does a good job of modeling the situation’s behavior.

r’s rpois function generates values from the Poisson distribution and returns the results. The function takes two arguments:

The number of observations you want to see

The estimated rate of events for the distribution; this is expressed as average events per period.

The expected syntax is :

rpois(# observations, rate = rate)

Continuing our example from above:

# r pois – poisson distribution in r examples

r pois(10, 10)

[1] 6 10 11 3 10 7 7 8 14 12

As you can see, there is some variation in the customer volume. A couple of minutes have seven or eight. One has 6. And apparently there was a mad dash of 14 customers as some point. The Poisson distribution models this type of variation in the expected throughput of a process.

Practical uses of Poisson distribution.

The Poisson distribution is commonly used within industry and the sciences.

The classical example of the Poisson distribution is the number of Prussion soldiers accidentally killed by horse-kick, due to being the first example of the Poisson distribution’s application to a real-world large data set. Ten army corps were observed over 20 years, for a total of 200 observations, and 122 soldiers were killed by horse-kick over that time period. The question is how many deaths would be expected over a period of a year, which turns out to be excellently modeled by the Poisson distribution.

Some other examples:

Number of equipment failures per day for logistics company

Number of customers arriving at a retailer

The number of visitors to a website

The number of inbound phone calls

Number of customer complaints