It Can't Happen That Often

The Cubs win the World Series. Bernie Madoff's crooked scheme crashes in a severe recession. These are rare events, and yet they have happened. Or not happened. With thirty major league teams, barring external conditions, one would expect the poor Cubbies to win the Series on average every thirty years. But it has been 102 years, although they won often in the 19th century. Oddly, though, that does not mean the sky is falling or aliens have invaded from space and don't like the Cubs.

"It's the law of averages; it's going to catch up with the other teams eventually." Actually, there is no such law in statistics, although there *is* well-known math that enables predicting these types of events, called "Poisson's Rain." (That's pronounced like "croissant" not like "poison.")

A good example of independent, randomly occurring, statistically regular events is falling raindrops. As far as anyone can tell, each raindrop falls on, say, a sidewalk block independently from all the other blocks. And there is, for reasonably short periods, say a few minutes, an average number per second. There many other events that exactly fit this model. For another example, how many calls will come into a customer-service phone number in a given time, if the average is known? (How many operators do you need to prevent a backlog from forming?) Our intuitions usually do not help us predict the actual number in a given time period, but the mathematicians have help for us.

There is a well established—albeit intimidating—formula that can tell us the probability (from zero to one, one being certain) that N raindrops will *actually* fall in a second if *on average (mean)* X fall in a second. (If you crave intimidation, go to "Poisson Distribution" on Wikipedia.) But Microsoft Excel includes a function for the Poisson probability, so you don't need to calculate it. Even easier, I have made a small spreadsheet that you can just plug into. Email me for a copy, Poisson at SmallBizLawyer.us.

For jollies, let's take an example. What if *each* of ten neighborhoods floods on *average* once every five years? Excel tells us the odds of no floods in a given year in one neighborhood are very high, 82%, while the chances of one flood go down to 16%—1/5 as likely. Two floods are less than 2% likely, and more than that not worth discussing.

While I was writing this essay, the Baltimore TV stations freaked out because four people in the city were murdered in one weekend. While I do not applaud such atrocious behavior by the shooters, the occurring of four in a weekend is to be expected in a city averaging 260 per year lately. On average 260 per year is about .71 per day *on average* (X = .71 per day) or 1.42 per 2-day weekend. The probability of exactly two in a weekend is .24, same as the odds of no murders. But the likelihood of having four in any random weekend is about 4%, not zero. That is about one of every 25 weeks. So you don't need to go along with the hype when the late news anchors scream about a deadly weekend.

Here is a practical business example that involves providing resources for peak demand. Let us say you hire enough technicians to handle an average of 100 service calls per day, and they are independent, statistically regular events. The (Poisson) odds are about 4% you will get exactly 100 in a given day. The results on a given day will be scattered, mostly numbers near 100 but higher or lower. If you graphed the results—the number of service calls along the horizontal axis, and the number of days of a specific number of calls vertically—you would have what statisticians call a “distribution curve.” Given enough days doing this, it would likely be bell shaped, similar to grades on tests.

The width of the distribution, defined in a special mathematical way, is called the “standard deviation” or S.D. For Poisson events, this is always the square root of the average number of events. We will look at a few numerical examples in a moment, to pin that idea down. About 2/3 of any Poisson distribution curve is between one S.D. below the average and one S.D. more than the average.

The square root of 100 you may recall (the number that when squared or multiplied by itself gives you 100) is ten. So, with our example of 100 service calls each day on average, on about 2/3 of the days you will get 100 plus or minus 10 calls. But on about 1 out of 6 days, you will get more than 110, and on one of 6 days less than 90. Under 90 shorts your hopes for revenue and over 110 probably overloads your shop’s ability to respond timely. How many technicians, then, should you hire? There is no right answer. You have to set customer expectations based on having personnel for the average or 100 service calls, recognizing that some days you will fall behind. Just because you got 120 today does not mean you will also get 120 tomorrow. Maybe it will be 80 and you can catch up. So far as anyone can tell, each day is independent of every other day.

Let us change the numbers and see what happens. With 25 service calls per day on average, the S.D. is the square root of 25 or 5. While with 100 calls an average day the S.D. is 10% of the average, with 25 (a smaller shop) the S.D. is 5/25 or 20% of the average. It is tougher to staff up for a small shop because the *peak load is statistically a bigger percent of the average load.* Put another way, one of 6 days will produce 30 or more calls, a real problem when you staff for 25. And one of six will have 20 or fewer calls, making revenue hard to keep up.

That is all the fault of mathematics, not bad management.

A quite big shop, averaging 400 calls a day, has an S.D. of the square root of 400 or 20. That is only 5% of the average (20/400). For this big shop, the likely range (2/3 of days) is as always plus or minus one S.D.—from 380 to 420 calls. One of six days will produce over 420, but 420 is not that much more than 400—they can be squeezed in.

Hits on a particular web page, by the way, are another case of Poisson events, subject to the same rules.

The results of having independent events that, nonetheless, follow statistical rules such as a measurable average number of events in a unit of time, are well predictable with the formulas attributed to Siméon Denis Poisson. Those results often are contrary to our intuitions, but math—as implemented now in Excel—can give more real results we can manage a business with.

Independently randomly occurring events that even so have statistical regularity, called Poisson events, have been studied in great depth the past 200 years. Formulas developed in that time can be applied to many practical business processes to predict what will happen and thus manage future conditions.