



CURRICULUM INSPIRATIONS: www.maa.org/ci



Uplifting Mathematics for All

www.theglobalmathproject.org



INNOVATIVE CURRICULUM ONLINE EXPERIENCES: www.gdaymath.com

TANTON TIDBITS: www.jamestanton.com



TANTON'S TAKE ON ...



THE POISSON DISTRIBUTION



DECEMBER 2017

I don't know statistics! I've only ever taken one—very basic—introductory course on the subject from which I feel I learned very little. The rest of my understanding of the subject has been from personal reading and from teaching my own course on the subject! But I've secretly been longing for understanding the history and theoretical underpinnings of the topic for a very long time: Why is the central limit theorem true – mathematically and intuitively? How does one actually figure out the chi-squared distribution? Why, really, do you divide by $n - 1$ in the formula for standard deviation? And so on. (I figured out one answer to that last question here: <http://www.jamestanton.com/wp-content/uploads/2012/03/Curriculum->

[Essay July-2014 MEAN-AND-VARIATION.pdf](#).)

I was reminded of my lack of knowledge of statistics last week at a university reception. An astronomer was chatting with his graduate student and said “Everything is the Poisson distribution” to then look at me, the mathematician, to say “Right?”

It might be shocking to learn that I don't know what the Poisson distribution is! Well, I didn't at that moment.

I have a maxim: *It is okay not to know. But it is not okay not to want to find out.*

So I decided this month to find out about the Poisson distribution and think about its mathematical meaning.

www.theglobalmathproject.org

www.jamestanton.com

www.gdaymath.com

In looking up the formula one retrieves:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \text{ for } k \in \{0, 1, 2, 3, \dots\}.$$

This reads: *Some random variable X can take on non-negative integer values and the chances that you'll actually see it adopt a particular value k is given by a crazy formula that depends on some constant parameter λ .*

Got it? Is it now clear that everything is Poisson?

Statisticians and mathematicians do not come up with formulas out of the air. What is the story behind this bizarreness?

PERHAPS EVERYTHING IS BINOMIAL?

If I roll a die three times, what are the chances I'll get exactly two sixes?

Maybe I'll first roll a non-six and then two sixes:

$$N \ 6 \ 6.$$

The chances of seeing this are $\frac{5}{6} \times \frac{1}{6} \times \frac{1}{6}$.

Or, following the implied notation, I could roll $6 \ N \ 6$ or $6 \ 6 \ N$ with

probabilities $\frac{1}{6} \times \frac{5}{6} \times \frac{1}{6}$ and $\frac{1}{6} \times \frac{1}{6} \times \frac{5}{6}$,

respectively.

So we see that there are 3 ways I could see two sixes and one non-six in a roll of three, each with the same probability. The chances of me seeing exactly two sixes is thus

$$3 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right).$$

In general, if in an experiment I have a probability p of seeing a success, and hence a probability $1 - p$ of seeing a failure, and I run the experiment n times

in a row, the probability that I will see exactly k successes is

$$\binom{n}{k} p^k (1-p)^{n-k}.$$

Here $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ is " n choose k ,"

the number of ways k successes can be arranged among a string n units long. For example, the chances of seeing exactly two sixes after rolling a die three times is

$$\frac{3!}{2!1!} \left(\frac{1}{6}\right)^2 \left(1 - \frac{1}{6}\right) = 3 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right),$$

as we saw.

We have here a *binomial distribution*. In general, for each success probability p and run length n we have the distribution

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k \in \{0, 1, 2, \dots, n\}$.

Here X is the random variable given as the number of times I see a success in running an experiment n times in a row."

In the 1700s scholars became interested in questions about runs of events that seem to occur in an ongoing fashion and counting the number of occurrences of the event you might expect to see in any selected time period. French mathematicians Siméon Denis Poisson and Abraham de Moivre were the first to develop the mathematics for this.

I've been sitting on my porch for many hours of my life, counting the number of cars that pass by each hour. After many hours of observation I can say that the flow of cars is more or less regular (there is no difference in traffic flow in the day versus

the night) and that, on average, 12 cars pass my house each hour.

I am about to go sit out on my porch for an hour. What is the probability I will see 15 cars?

One possible start to this question is to treat it as a run of an experiment, not over an hour but, instead, one experiment per minute for 60 minutes. If I know the chances p of seeing a car in a given minute, then the chances of seeing 15 cars over a course of 60 minutes would be

$$\binom{60}{15} p^{15} (1-p)^{45}.$$

How might I estimate p ?

Well, I expect to see an average of 12 cars over the course an hour. So, on average, I expect to see one car in any five-minute period. So chances of me selecting a minute with the appearance of a car in it should be

$$p = \frac{1}{5}.$$

That's it. We have a formula for the chances of seeing 15 cars over the course of an hour.

$$\binom{60}{15} 0.2^{15} 0.8^{45} \approx 0.076$$

There is a problem with our work here: we've assumed that cars are spaced apart so that I'll never see two cars within the same minute. To obviate this objection, let's work instead with the 3600 seconds in an hour: I've never seen two cars go by within the same second.

Now, on average, over the 3600 seconds of an hour, 12 of them will have a car "in them," and so the chances of me seeing a

car in any particular second is $p = \frac{12}{3600}$.

This now gives

$$\binom{3600}{15} \left(\frac{12}{3600}\right)^{15} \left(1 - \frac{12}{3600}\right)^{3585} \approx 0.080$$

for the chances of me seeing exactly 15 cars.

Well, actually, I realise now that I think I have seen two cars go by within the same second: two cars going in opposite directions passed each other in front of my house.

So maybe I should repeat this analysis not over every minute or every second, but instead over every nanosecond or every picosecond. I need a very short length of time I can be sure will never have two cars appearing in it.

Let's break the hour into n , almost instantaneous, units of time. The chances of me seeing a car within any one of those

units is $p = \frac{12}{n}$ and so the probability of

me seeing 15 cars over a run of n of these units of time is

$$\binom{n}{15} \left(\frac{12}{n}\right)^{15} \left(1 - \frac{12}{n}\right)^{n-15}.$$

Ideally, we should compute this for larger and larger values of n . That is, we should take the limit as n grows infinitely large.

⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡⚡ Taking the Limit

The formula we have reads

$$\frac{n \cdot (n-1) \cdots (n-14)}{15!} \cdot \frac{12^{15}}{n^{15}} \cdot \left(1 - \frac{12}{n}\right)^n \cdot \frac{1}{\left(1 - \frac{12}{n}\right)^{15}}$$

I can see how parts of this formula change as n grows larger and larger.

For instance

$$\frac{1}{\left(1 - \frac{12}{n}\right)^{15}} \rightarrow \frac{1}{(1-0)^{15}} = 1$$

as n grows.

Also,

$$\frac{n \cdot (n-1) \cdot (n-2) \cdots (n-14)}{n^{15}} = \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{14}{n}\right) \rightarrow 1 \cdot 1 \cdots 1 = 1$$

as n grows.

And the formula $\left(1 - \frac{12}{n}\right)^n$ is reminiscent of the famous compound-interest formula

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e,$$

or more generally,

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x.$$

So we have that

$$\left(1 - \frac{12}{n}\right)^n \rightarrow e^{-12} \text{ as } n \text{ grows.}$$

So we see that our formula,

$$\binom{n}{15} \left(\frac{12}{n}\right)^{15} \left(1 - \frac{12}{n}\right)^{n-15},$$

in the absolute ideal of looking at finer and finer intervals over the hour, becomes the formula

$$1 \cdot \frac{12^{15}}{15!} \cdot e^{-12} \cdot 1 = \frac{12^{15}}{15!} e^{-12}.$$

This has value ≈ 0.072 . There is about a 7% chance I will see 15 cars over any given hour.

Aside: From $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$ we

see that

$$\left(1 + \frac{x}{n}\right)^n = \left(\left(1 + \frac{1}{n/x}\right)^{n/x}\right)^x \rightarrow (e)^x$$

if x is positive (since n/x grows large and positive if n does).

We also establish

$$\lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right)^n = e^{-1}$$

by observing that

$$\begin{aligned} \left(1 - \frac{1}{n}\right)^n &= \left(\frac{n-1}{n}\right)^n = \frac{1}{\left(\frac{n}{n-1}\right)^n} \\ &= \frac{1}{\left(1 + \frac{1}{n-1}\right)^{n-1} \cdot \left(1 + \frac{1}{n-1}\right)} \end{aligned}$$

$$\rightarrow \frac{1}{e \cdot (1+0)} = \frac{1}{e}$$

as n grows.

Now we can follow similar algebra as above to show that

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x$$

even if x is negative.

Most important, this work has led us to the Poisson probability distribution formula!

If you have a phenomenon that seems to occur at a more-or-less steady rate, with the number occurrences during any set time length seeming to be more-or-less the same over all periods of that length, and no two events can occur at exactly the same instant, then, if you have good reason to believe that, for a given period length, on average λ events occur, the probability that you will see k events during any particular time period of that length is

$$\frac{\lambda^k}{k!} e^{-\lambda}.$$

This is the Poisson distribution.

I am guessing that many astronomical phenomena occur at more-or-less steady rates, on average—meteor strikes to the Earth, appearance of Sun spots—and so maybe close to everything in astronomy is indeed Poisson?

