Some very brief notes on the Poisson Distribution

Suppose a certain event happens, on average, λ times on each interval of a given extent.

Examples:

1. On average, in a meteor shower, five meteorites hit the earth every square kilometer.
2. On average, on a certain block in Times Square, five people spit out or throw their gum on the ground every day.
3. On average, in the Boulder Target during peak hours, 25 people enter the self-checkout line every 15 minutes.
4. On average, a computer CPU running certain software receives three instructions per nanosecond.
5. On average, a sample of polonium emits 3.8715 α-rays per 7.5 second period.
6. On average, a large computer program contains 4.3 errors per 10,000 lines of code.

Let X denote the number of actual occurrences of this event per interval of the given extent. Question: what is the probability mass function for X? That is, given any integer k with $k \geq 0$, what is $P(X = k)$?

(Remark: in contrast to the binomial distribution, in the present case, there is no upper bound for k. Here, k can, in theory, be as large as we like.)

To answer our question, let's choose a large integer N, and divide our original interval up into N subintervals of equal length. Then, since our event happens λ times, on average, on our original, larger interval, then we would expect that it happens, on average, λ/N times on each of the smaller subintervals. That is, the expected number of occurrences of our event on each subinterval is λ/N.

Suppose we've chosen N large enough that the event is very unlikely to happen *more than once* on any subinterval. Then, essentially, the event happens either once or zero times on a subinterval. That is, the number of times the event happens on a subinterval is a binomial random variable, with expected value λ/N. For such a variable, as we've seen, the probability of the event occurring equals its expected value. So the *probability* of the event happening on a given subinterval is λ/N.

For the event to happen k times in total, on the original, larger interval, it must happen on exactly k of the N subintervals. So we have, in essence, N trials of a binomial random variable with $P(\text{success}) = \lambda/N$. We know that, in such a situation, the probability of k successes is

$$P(X = k) \approx \binom{N}{k} \left(\frac{\lambda}{N}\right)^k \left(1 - \frac{\lambda}{N}\right)^{N-k}.$$

This is only approximately true, because it's only approximately true that our event can happen at once most per subinterval. But now we note: in the above, we took N to be any large number. And it stands to reason that, the larger N is, the better an approximation we get. In sum, we have the following.

CONCLUSION. Suppose a certain event occurs, on average, λ times in each interval of a given extent. Let X be the actual number of times it happens in a given interval of such an extent. Then

$$P(X = k) = \lim_{N \to \infty} \binom{N}{k} \left(\frac{\lambda}{N}\right)^k \left(1 - \frac{\lambda}{N}\right)^{N-k}.$$

The above formula is somewhat cumbersome: who wants to have to take limits all the time? Fortunately, using some results from calculus (for example, l,Hôpital's rule), one can deduce, directly from the above CONCLUSION, the following fairly simple formula.

Theorem 1. Under the conditions of the above CONCLUSION, we have

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

for $k = 0, 1, 2, 3, \ldots$.

In the situation described above, we say "X has a Poisson distribution, with parameter λ."

Example. In 1911, Ernest Rutherford et al. determined that, on average, a sample of polonium emitted 3.8715 α-rays every 7.5 seconds. Find the probability that, in a given 7.5 second interval, the sample emits 2, 3, or 4 α-rays.

Solution. The number X of α-rays emitted every 7.5 seconds has a Poisson distribution with parameter $\lambda = 3.8715$. So, by Theorem 1,

$$P(X = 2) = \frac{3.8715^2}{2!} e^{-3.8715} = 0.156084,$$

and similarly $P(X = 3) = 0.201426$, $P(X = 4) = 0.194955$.

Finally we describe the mean, variance, and standard deviation of a Poisson random variable.

Theorem 2. If X is Poisson with parameter λ, then

$$E(X) = \text{var}(X) = \lambda \quad \text{and} \quad \text{std}(X) = \sqrt{\lambda}.$$

Proof. We prove the formula for $E(X)$ only; the formula for $\text{var}(X)$ is deduced similarly (though the details are somewhat messier).

We have

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} k \cdot P(X = k) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} \\ &= 0 \cdot \frac{\lambda^0}{0!} e^{-\lambda} + 1 \cdot \frac{\lambda^1}{1!} e^{-\lambda} + 2 \cdot \frac{\lambda^2}{2!} e^{-\lambda} + 3 \cdot \frac{\lambda^3}{3!} e^{-\lambda} + 4 \cdot \frac{\lambda^3}{4!} e^{-\lambda} + \cdots \\ &= \lambda e^{-\lambda} + \lambda^2 e^{-\lambda} + \frac{\lambda^3}{2} e^{-\lambda} + \frac{\lambda^4}{3!} e^{-\lambda} + \cdots = \lambda e^{-\lambda} \left(1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{3!} + \cdots\right) \\ &= \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda. \end{aligned}$$

The next-to-last step is by the power series expansion for e^{λ}.