**Poisson distribution—a thumbnail description**

**Trials, opportunities and events: the premises of the Poisson distribution**

We’ll define a “trial” generically as a collection of a large number *N* of “opportunities.” *N* is very large and is identical for all trials. For example, a trial could be a sample of 6 million yeast cells, or a sample of 500 products in an manufacturing quality control (QC) study.

An “opportunity” is an opportunity for an “event” to occur. For example, the event could be that a yeast cell is a mutant of some specified type, or that a product in a QC study is defective.

The probability *p* for an event to occur in an individual opportunity is very small, and is the same for all opportunities. The probability that an event occurs in a given opportunity is independent of the occurrence or non-occurrence of events in any other opportunity.

The expected number of events *µ* in a trial is just the probability of an event per opportunity (*p*) times the number opportunities per trial (*N*): *µ* = *p*×*N*. Even though *p* is very small, because *N* is very large *µ* need not be small (though of course it’s much smaller than *N*).

**Poisson distribution**

Given the above premises, it’s not hard to show (using elementary calculus) that Pr(*m*|*µ*), the probability that exactly *m* events will occur in a trial given that the expected number of events is *µ*, is given by the Poisson distribution function:, for *m* = 0, 1, 2, 3,…, where *e* is the base of the natural logarithms and *m*! = “*m* factorial” = *m*(*m*–1)(*m*–2)…(2)(1). By convention, 0! = 1. In particular, , where in the last equation we’re using the functional notation for . We won’t try to prove that the Poisson distribution function is valid under the assumptions of the previous section.

Like all probability distributions, the sum of the Poisson probabilities for all possible values of *m* from 0 to infinity is 1: ; we won’t try to prove that either.

**Expected number of events per trial and expected variance in the number of events per trial according to the Poisson distribution**

If we analyzed a huge number of trials obeying the premises of the Poisson distribution in the first section, the mean (average) number of events per trial, , would be *µ*. The variance in the number of events per trial—that is, the value of  averaged over a huge number of trials—would also be *µ*. These average values of statistics for huge numbers of trials are called the statistics’ *expected* values. So the expected value of both *m* and —written  and , respectively—is *µ* for trials obeying the premises of the Poisson distribution. (It’s not hard to prove these statements, using only high school algebra; no calculus is required.)

If we analyze a more modest number of trials, the mean and variance of *m* won’t be exactly the same, but they’ll be pretty close. Conversely, one way to check whether or not a series of trials is obeying the Poisson distribution is to see if the mean and variance in the event numbers *m* are pretty close to each other. In the yeast fluctuation test, for example, the post-exposure hypothesis predicts that the red colony counts on the individual culture dishes, like the counts on the bulk culture sample dishes, will be distributed according to the Poisson distribution. If it turns out that the variance is much higher than the mean, that would be evidence against the post-exposure hypothesis.