A Novel Use of the MicroSet.

Academic physics technicians and demonstrators are fond of saying they use a device in an unintended way. I'm rather certain Mr. Mumford never expected someone would use his device to collect statistical data in an attempt to show that popping corn pops are Poisson distributed, and to give an example of the well known "truth" of that distribution for radioactive decay.

The MicroSet is a nearly ideal device used with its interface software to collect statistical Poisson and Poisson Interval distribution data. I'll devote a short space to a description of these distributions before discussing popping corn pops and Geiger Counter data, Those already familiar should skip this section.

Poisson Probabilities

The Wikipedia entry is a good description of this probability distribution. It begins with:
"In probability theory and statistics, the **Poisson distribution** (pronounced [pwasɔ̃]) is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event. The Poisson distribution can also be used for the number of events in other specified intervals such as distance, area or volume.

“For instance, suppose someone typically gets 4 pieces of mail per day on average. There will be, however, a certain spread: sometimes a little more, sometimes a little less, once in a while nothing at all. Given only the average rate, for a certain period of observation (pieces of mail per day, phone calls per hour, etc.), and assuming that the process, or mix of processes, that produces the event flow is essentially random, the Poisson distribution specifies how likely it is that the count will be 3, or 5, or 10, or any other number, during one period of observation. That is, it predicts the degree of spread around a known average rate of occurrence.”

One finds the average rate of events by dividing the total number over many days. Using their mail example, with the average number four per day, what is the probability of some other number? For example P(0), no mail or P(6) six pieces? The answer is:

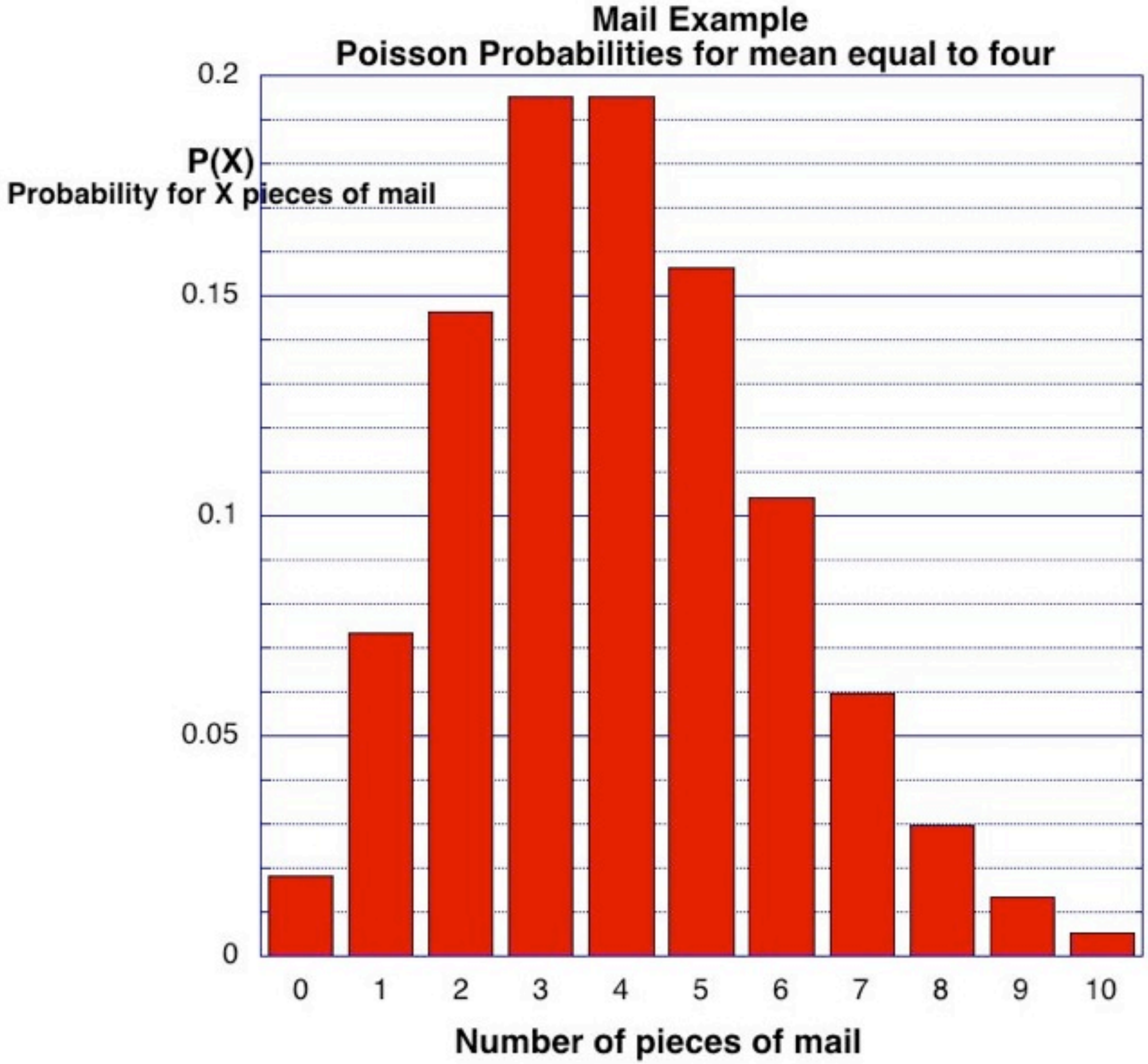
Equation one:

$$P_X = \frac{m^X}{X!} e^{-M}$$

Where x is the number of pieces and m is the average (4). ! Is the symbol for factorial. [x times x-1 times x-2 ... all positive. Note: !0 = 1]. *e* is Euler’s number, approximately 2.718.

Here are the first eleven probabilities:

P(0) ~= 0.0183
P(1) ~= 0.0733
P(2) ~= 0.1465
P(3) ~= 0.19537
P(4) ~= 0.19537
P(5) ~= 0.1563
P(6) ~= 0.1042
P(7) ~= 0.0595
P(8) ~= 0.0298
P(9) ~= 0.0132
P(10) ~= 0.00529



Note the asymmetry, and, because it is a discrete distribution, I've used a bar chart.

The most familiar distribution is the symmetrical bell shaped normal or Gaussian. For large means (greater than 100) the Poisson approximates the Gaussian, and in the limit m => infinite, they are the same. Finally, obviously, the sum of all of the probabilities (for any distribution) must be equal to one. In this case the first 11 equals ~ 0.997

The Poisson Interval distribution

"The interval distribution is derived from Poisson's distribution and describes the distribution in size of the of the time intervals between successive events in any random process in which the mean rate has the constant value of '*m'* events per unit time." [p. 753 ***The Atomic Nucleus****,* Robley D. Evans, 1955]

The probability there will be no events (pieces of mail, for example) for a time interval *t*, during which there should be *mt* events on average is (our example: one day and mean, four/day) substituting in equation one:

$$P_0 = \frac{(mt)^0}{0!} e^{-mt} = e^{-mt}$$

And the probability that there will be an event in the time interval *dt* is simply *m dt.* Therefore, the combined probability there will be no event during the time interval *t*, and one event in the following time *dt,* (*t* + *dt*), is the product, *i.e.*, *mdt* e^{-mt}, or $me^{-mt}dt$. Continuing, to quote Dr. Evans: “Hence, in a random distribution which follows the Poisson distribution and has a constant average interval 1/m, the probability *dP*(*t*) that the duration of of a particular interval will be between *t* and *t* + *dt*

is $$dP_t = me^{-mt}dt$$ Equation two.

“We see at once that *small time intervals have a higher probability* than large time intervals between the randomly distributed events.” If the data includes a large number of, *N*, intervals, then the number of intervals

greater than t_1, but less than t_2 is $n = N\int_{t_1}^{t_2} me^{-mt}\,dt$, or $n = N(e^{-mt_1} - e^{-mt_2})$, eq. (3).

The MicroSet

I found two modes useful for collecting Poisson data, Time 1 and Dark / Light. Time one is useful for small means, less than one event per second, and Dark / Light for larger ones. This is because the minimum time between recordable events for Time 1 is approximately 75ms, and the maximum is approximately 83s. Time 1 measures the time between the leading edge of a pulse and the leading edge of the next pulse within the resolution just given. Another requirement is the pulses' width must be greater than about six microseconds (both modes), and, of course less than the time between succeeding pairs. More accurately, somewhat less than just overlapping. *e.g.* For a pulse separation of 47ms the width must be less than one ms. This specification accommodates the background radiation detected by a small Geiger-Müller tube. *i.e.* Approximately 0.2 counts (events) per second. On the other hand, Dark / Light's minimum time between pulses is approximately 23 ms and maximum of ten seconds. This approximately accommodates domesticate popping corn. One of my better popping corn trials yielded, in over 75 seconds, 398 captured pops, or approximately 5.3 pops per second. Using eq. (3) one finds an approximate 12% loss. Nearly all of which due to the 23ms resolution. [$t1$ = 0.023, and $t2$ = 10s]. The case for G-M tube background radiation is somewhat better. In a trial where the tube is shielded by several inches of lead, over the time of 16,417s 2816 counts were recorded, which is 0.1715 cps. Again using eq (3) there is a less than 2% loss. Certainly ignorable. However, the 12% loss popping is more serious, because the measured mean is less than the actual one resulting in the calculated loss being too little. An approximate correction may be made by iteration.

Data Analyses and Results

To compare the data to the expected Poisson distributions one bins the data according to numbers of events in time intervals (in our two examples). It's generally agreed that to use the most common "Goodness of Fit" test, the Chi square, one must have a minimum of five groups (bins) with a minimum of five points in each. Hence a required minimum of 25 points. However, depending on the mean many more points /