Derivation of the Poisson distribution (the Law of Rare Events).

Phys353 lecture note additions
Jim Remington, Dept. of Physics, University of Oregon

We begin with the exact result for the probability distribution governing the outcome of N tosses of a very unfair coin. Assume that p(heads), the probability of obtaining heads on one toss, is much less than one **(p<<1)**. Thus, in a long sequence of coin tosses, the appearance of heads is a rare event. As shown in class, (independent of the value of p), the probability of observing n heads in N tosses is:

$$P(N,n) = \frac{N!}{(N-n)!n!} p^n (1-p)^{N-n}$$

We make two approximations:
First show that:

$$(1-p)^{N-n} \approx e^{-Np}$$

Derivation:

$$\ln[(1-p)^{N-n}] = (N-n)\ln(1-p)$$
$$\ln(1-p) \approx -p \quad (for\ \ p << 1)$$
$$(N-n)(-p) \approx -Np$$
$$therefore$$
$$(1-p)^{N-n} \approx e^{-Np}$$

Second, show that:

$$\frac{N!}{(N-n)!} \approx N^n$$

Derivation:
Using Stirlings approximation

$$\ln(N!/(N-n)!) \approx N\ln N - N - (N-n)\ln(N-n) + (N-n)$$
$$n << N$$
$$\ln(N-n) = \ln N + \ln(1 - n/N) \approx \ln N - n/N$$
$$(N-n)(\ln N - n/N) = N\ln N - n\ln N - n + n^2/N$$
$$n^2/N \approx 0$$
$$so$$
$$\ln(N!/(N-n)!) \approx n\ln N$$
$$thus$$
$$N!/(N-n)! \approx N^n$$

(You can see that this approximation makes sense by trying some values, for example 100!/98! ~ 100^2)

Putting these results together

$$P(N,n) \approx (Np)^n e^{-Np} / n!$$

Often, the symbol **λ = Np** is substituted, where λ is the expected outcome for a given sample interval.

In the coin-toss example, λ would the number of heads expected in N tosses, given the probability p that heads will come up in one toss. To put it the other way around, if you wished to determine p, you would toss the coins N times, count λ heads, and calculate p ~ λ/N.

So to summarize,

$$P(N,n) = \lambda^n e^{-\lambda} / n!$$

P(N,n) is the Poisson distribution, an approximation giving the probability of obtaining exactly n heads in N tosses of a coin, where (p = λ/N) <<1.

To think about how this might apply to a sequence in space or time, imagine tossing a coin that has p=0.01, 1000 times. This will produce a long sequence of tails but occasionally a head will turn up.

TTTTTTTTTTTTTHTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTHTTTTT…

(about 10 heads are expected total). Now consider this sequence to be seconds or minutes in time, where H means that a customer arrives at a cash register stand and T means no customer. The customer arrival rate λ is 10/1000 seconds, or 1 per 100 seconds. The Poisson distribution then gives us, choosing for example 100 second intervals, the probability that n = 0, 1, 2, … customers arrive in any given 100 second interval. The result can be extended to any desired time interval by choosing λ appropriately.

Alternatively, one could consider the probability of encountering road kill per mile or kilometer. H would correspond to road kill observed in a particular distance interval, T=none.

Suppose that from 10-11 am, 60 customers arrive at the bank, or 1 customer per minute. On a per minute basis we can calculate the probability that exactly 0,1,2,… customers will arrive:

$$P(n) = 1^n e^{-1} / n!$$
$$P(0) = 1/e = 0.368$$
$$P(1) = 1/e = 0.368$$
$$P(2) = e^{-1} / 2! = 0.184$$
$$\ldots$$

To calculate the probability that 1 or more customers will arrive in a given minute, you sum (integrate) the distribution, i.e. P(1) + P(2) + P(3) + … = 1 - P(0) = 0.632

For one nice discussion of the properties of the Poisson distribution, see
http://en.wikipedia.org/wiki/Poisson_distribution

Queuing Theory

An interesting and important application of the Poisson distribution is in Queuing Theory (supermarket cashier lines, freeway traffic density considering on/off ramps, etc.), for an introduction see http://www.eventhelix.com/RealtimeMantra/CongestionControl/queueing_theory.htm

The story begins with the simple but surprisingly non-obvious theorem known as Little's Theorem. The following section in small size font is taken from the above web site.

The average number of customers (N) can be determined from the following equation:

$$N = \lambda T$$

Here lambda is the average customer arrival rate and T is the average service time for a customer. Proof of this theorem can be obtained from any standard textbook on queueing theory. Here we will focus on an intuitive understanding of the result. Consider the example of a restaurant where the customer arrival rate (lambda) doubles but the customers still spend the same amount of time in the restaurant (T). This will double the number of customers in the restaurant (N). By the same logic if the customer arrival rate remains the same but the customers service time doubles, this will also double the total number of customers in the restaurant.

Note: the above argument is only a rough guide, because if customers arrive more rapidly than they can be serviced, N will continue to increase without bound. After some point, it gets exciting!

To briefly explore the consequences of Little's Theorem, let's go back to the bank example, where we saw that on average, one customer arrives per minute. Supposing that it takes 10 minutes to service the customer, Little's Theorem says that on average, 10 customers will be in the bank.

It turns out that 10 is also the minimum number of bank tellers needed!

As a check, note that one bank teller can service on average 6 customers/hour (1 in 10 minutes), but 60 will have arrived in that hour, so 10 tellers are needed to keep everyone happy. The quandary is this: banks don't want to hire too many tellers but on the other hand, customers get upset if they have to wait in long lines. The number of customers fluctuates, so tellers will sometimes have to do other jobs.