The Poisson Distribution.

A) <u>Example:</u> A computer program contains, on average, 4 errors per 10,000 lines of code. Let X be the actual number of errors found in a 10,000-line block. Find $P(X=k)$ $(k \geq 0)$.

<u>Solution</u>

In a <u>single</u> line of code, only <u>rarely</u> would we expect <u>more</u> than one error. So in a single line, we'd expect a "success" (error) or "failure" (no error), with $P(\text{success}) = \frac{4}{10000}$.

So X, the number of errors in 10,000 lines, is approximately $B(10{,}000, \frac{4}{10{,}000})$. (assuming independence of errors). So

$$P(X=k) \approx \binom{10{,}000}{k} \left(\frac{4}{10{,}000}\right)^k \left(1 - \frac{4}{10{,}000}\right)^{10{,}000-k}. \qquad (*)$$

For example,

$$P(X=6) \approx 0.104206.$$

Now if we break up our 10,000 lines into 20,000 pieces, instead of 10,000, then we get $(*)$ with 20,000 instead of 10,000. Moreover, our approximation <u>improves.</u>

In general, replacing 10,000 with N, above, gives a better approximation the larger N is. Conclusion:

$$P(X=k) = \lim_{N\to\infty} \binom{N}{k} \left(\frac{4}{N}\right)^k \left(1-\frac{4}{N}\right)^{N-k}.$$

COOL FACT (see HW #8): this limit equals

$$\frac{4^k}{4!} e^{-4} \quad !!$$

For example, the probability of 6 errors in a block of 10,000 lines is

$$P(X=6) = \frac{4^6}{6!} e^{-4} = 0.104196.$$

B) Generalization: the Poisson distribution.

Consider an event that happens, on average, λ times in every interval (of time, space, etc.) of some fixed extent.

Examples.

(a) On average, in a meteor shower, 5 meteorites hit the earth every square kilometer.

(b) On average, in Target during peak hours, 25 people enter the self-checkout line every 15 minutes.

(c) On average a computer CPU, running certain software, receives 3 instructions per nanosecond.

(d) On average a sample of polonium emits 3.8715 α-rays per 7.5 second period.

Let X be the number of times the event actually occurs in an interval of the given extent. Then X is said to be a Poisson rv with parameter λ (in short: X is $P(\lambda)$), and

$$\boxed{P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (k=0,1,2,3,\dots)}$$

pmf for a Poisson random variable

Example 1.

In 1911, Ernest Rutherford et al. observed that, on average (over several hours), a sample of polonium emitted $\lambda = 3.8715$ α-rays every 7.5 seconds.

Let X be the number of rays emitted in a random 7.5 second period. Find $P(X=k)$ for $k \geq 0$.

Solution.

We have

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda} = \frac{3.8715^k e^{-3.8715}}{k!}.$$

E.g.

$$P(X=0) = \frac{3.8715^0 e^{-3.8715}}{0!} = e^{-3.8715} = 0.0208.$$

Here are the first few values of the pmf:

k	$P(X=k) = 3.8715^k \cdot e^{-3.8715}/k!$
0	0.0208
1	0.0806
2	0.1561
3	0.2014
4	0.1950
5	0.1510

Next time: mean and variance of a Poisson r.v.