Consider a Bernoulli process with some large number N of trials, and some small probability p of success on each trial. In the limit of N very large and p small, the probability $p(m)$ of m successes turns out to be parametrized entirely by just the mean number of successes, $z = Np$. First recall that

$$p(m) = \binom{N}{m} p^m (1-p)^{N-m} \ .$$

For $N \gg m$, we can approximate $N!/(N-m)! = N(N-1)\cdots(N-m+1) \approx N^m$, so $\binom{N}{m} = \frac{N!}{m!(N-m)!} \approx \frac{N^m}{m!}$, and

$$p(m) \approx \frac{1}{m!} N^m (z/N)^m (1-z/N)^{N-m} \approx \frac{z^m}{m!} \lim_{N\to\infty} (1-z/N)^N = \mathrm{e}^{-z} \frac{z^m}{m!} \ ,$$

where the factor of $(1-z/N)^{-m}$ can be ignored since by assumption $N \gg z$.

Note that the N dependence drops out of this probability in the limit as $N \to \infty$, with average z fixed (so that $p \to 0$). The form $p(m) = \mathrm{e}^{-z} \frac{z^m}{m!}$ is known as a Poisson distribution. (Note also that $\sum_{m=0}^{\infty} p(m) = \mathrm{e}^{-z} \sum_{m=0}^{\infty} \frac{z^m}{m!} = \mathrm{e}^{-z} \cdot \mathrm{e}^{z} = 1$, so these are properly normalized probabilities.)

We can also calculate the mean and variance of the distribution. For the mean,

$$\begin{aligned} E[m] &= \sum_{m=0}^{\infty} m p(m) = \sum_{m=0}^{\infty} m \frac{z^m}{m!} \mathrm{e}^{-z} = \mathrm{e}^{-z} \sum_{m=0}^{\infty} z \frac{\partial}{\partial z} \frac{z^m}{m!} \\ &= \mathrm{e}^{-z} z \frac{\partial}{\partial z} \sum_{m=0}^{\infty} \frac{z^m}{m!} = \mathrm{e}^{-z} z \frac{\partial}{\partial z} \mathrm{e}^{z} = \mathrm{e}^{-z} z \mathrm{e}^{z} = z \ . \end{aligned}$$

Recall that $\mathrm{Var}[m] = E\big[(m - E[m])^2\big] = E[m^2] - (E[m])^2$, where

$$\begin{aligned} E[m^2] &= \sum_{m=0}^{\infty} m^2 p(m) = \sum_{m=0}^{\infty} m^2 \frac{z^m}{m!} \mathrm{e}^{-z} = \mathrm{e}^{-z} \sum_{m=0}^{\infty} \Big(z \frac{\partial}{\partial z}\Big)^2 \frac{z^m}{m!} \\ &= \mathrm{e}^{-z} z \Big(z \frac{\partial}{\partial z}\Big)^2 \sum_{m=0}^{\infty} \frac{z^m}{m!} = \mathrm{e}^{-z} \Big(z \frac{\partial}{\partial z}\Big)^2 \mathrm{e}^{z} = \mathrm{e}^{-z} (z + z^2) \mathrm{e}^{z} = z + z^2 \ . \end{aligned}$$

So $\mathrm{Var}[m] = E[m^2] - (E[m])^2 = z + z^2 - z^2 = z$, and $\mathrm{Std}[m] = \sqrt{z}$. So the mean equals the variance for a Poission distribution, and the relative error goes as $\mathrm{Std}[m]/\mathrm{E}[m] = 1/\sqrt{z}$.

According to the current version (1 Oct 2013) of the Wikipedia entry,
http://en.wikipedia.org/wiki/Poisson_distribution :

"Applications of the Poisson distribution can be found in many fields related to counting:

- Telecommunication example: telephone calls arriving in a system.
- Astronomy example: photons arriving at a telescope.
- Biology example: the number of mutations on a strand of DNA per unit length.
- Management example: customers arriving at a counter or call centre.
- Civil engineering example: cars arriving at a traffic light.
- Finance / insurance example: number of Losses/Claims occurring in given time period
- Earthquake seismology example: seismic risk for large earthquakes.
- Radioactivity example: Decay of a radioactive nucleus.

Examples of events that may be modelled as a Poisson distribution include:

- The number of soldiers killed by horse-kicks each year in each corps in the Prussian cavalry. This example was made famous by a book of Ladislaus Josephovich Bortkiewicz (1868–1931).
- The number of yeast cells used when brewing Guinness beer (made famous by William Sealy Gosset 1876–1937).
- The number of phone calls arriving at a call centre within a minute.
- The number of goals in sports involving two competing teams.
- The number of deaths per year in a given age group.
- The number of jumps in a stock price in a given time interval.
- The number of times a web server is accessed per minute.
- The number of mutations in a given stretch of DNA after a certain amount of radiation.
- The proportion of cells that will be infected at a given multiplicity of infection.
- The arrival of photons on a pixel circuit at given illumination over given time period.
- The targeting of V-1 flying bombs on London during World War II."

A previous version (24 Oct 2006) also had

- "The number of cars that pass through a certain point on a road during a given period of time.
- The number of spelling mistakes a secretary makes while typing a single page.
- The number of edits per hour recorded on Wikipedia's Recent Changes page
- The number of roadkill found per unit length of road.
- The number of unstable nuclei that decay within a given period of time in a piece of radioactive substance.
- The number of pine trees per unit area of mixed forest.
- The number of stars in a given volume of space.
- The distribution of visual receptor cells in the retina of the human eye.
- The number of V2 rocket attacks per area in England, according to the fictionalized account in Thomas Pynchon's Gravity's Rainbow.
- The number of light bulbs that burn out in a certain amount of time.
- The number of viruses that can infect a cell in cell culture."

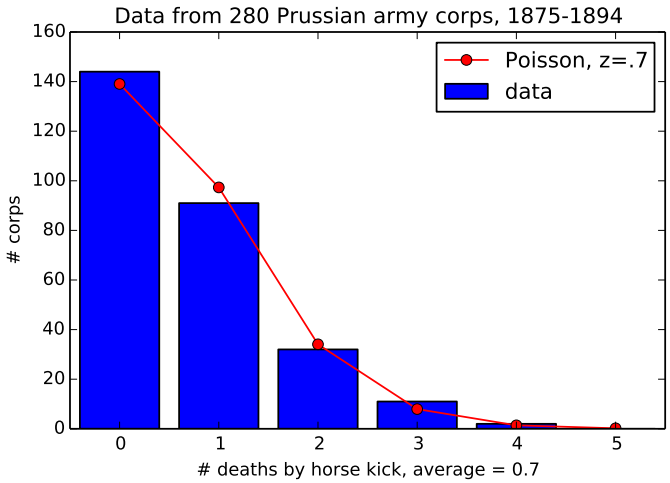
A classic example is of soldiers killed by horse-kicks in class, based on data collected from the Prussian army from 1875–1894. The army was divided into 280 corps, each containing a large number of soldiers, and data was obtained on the average number of deaths due to horse-kicks (a relatively rare event) per corp:

#deaths/year	#corps
0	144
1	91
2	32
3	11
4	2
≥ 5	0

From this data, the total average number of deaths/year is $0\cdot 144+1\cdot 91+2\cdot 32+3\cdot 11+4\cdot 2 = 196$, so the average number of deaths/year per corp is $196/280 = .7$. According to a Poisson distribution, the probability that a corp will have m deaths/year is thus $p(m) = \frac{(.7)^m}{m!}\mathrm{e}^{-.7}$:

#deaths/year	data	$\frac{(.7)^m}{m!}\mathrm{e}^{-.7}$
0	144/280=.51	.5
1	91/280=.33	.35
2	32/280=.11	.12
3	11/280=.04	.03
4	2/280=.01	.005
≥ 5	0/280=0	.0007

The probabilities predicted by the Poisson distribution are in close agreement with the data. Note that one number, the average number of deaths/year per corp of .7, permits understanding all six of the data points above. Note also that the corps with fewer or greater than average deaths are not somehow responsible for implementing safer or less safe practices, since the numbers follow the expected statistical distribution. (Only if they deviated from the expected Poisson statistics would additional explanation be necessary.)




```
corps=[144, 91, 32, 11, 2, 0]
z=sum([j*corps[j] for j in range(6)])/float(sum(corps))
poisson=array([exp(-z)*z**m/factorial(m) for m in range(6)])

bar([0,1,2,3,4,5], corps, align='center', label='data')
plot(sum(corps)*poisson, 'ro-', label='Poisson, z='+str(z))
xlabel('# deaths by horse kick, average = '+str(z))
ylabel('# corps')
title('Data from 280 Prussian army corps, 1875-1894')
xlim(-.5,5.5)
legend(numpoints=1)
savefig('horsekicks.pdf')
```

Another example is given by the number of distinct hostnames hitting a web site in any given second. Here the number of potential users is large, but there's only a small probability that any given one will be active within a one second period. The arXiv.org data below are taken from two different one-hour periods on 24 Oct 2006: one a relatively quiet hour starting at midnight, during which there was an average of $(0 \cdot 717 + 1 \cdot 1156 + 2 \cdot 922 + 3 \cdot 493 + 4 \cdot 232 + 5 \cdot 63 + 6 \cdot 12 + 7 \cdot 3 + 8 \cdot 2)/3600 = 5831/3600 = 1.62$ distinct hosts/second; the second a more active hour from 10:00–11:00 in the morning with an average of $(0 \cdot 106 + 1 \cdot 414 + 2 \cdot 714 + 3 \cdot 767 + 4 \cdot 674 + 5 \cdot 455 + 6 \cdot 268 + 7 \cdot 131 + 8 \cdot 46 + 9 \cdot 16 + 10 \cdot 6 + 11 \cdot 2 + 12 \cdot 1)/3600 = 12245/3600 = 3.40$ distinct hosts/second. The data are prescreened to eliminate robotic activity, and are again well-described by Poisson distributions.

Data for 24 Oct 2006, 00:00–01:00 EDT

#hosts/sec	data	$\frac{(1.62)^m}{m!}e^{-1.62}$
0	$717/3600 = .199$	.198
1	$1156/3600 = .321$	.320
2	$922/3600 = .256$	.260
3	$493/3600 = .137$	.140
4	$232/3600 = .064$	.057
5	$63/3600 = .0175$	.0184
6	$12/3600 = .0033$	.0050
7	$3/3600 = .0008$	.0011
8	$2/3600 = .0006$	.0002

Data for 24 Oct 2006, 10:00–11:00 EDT

#hosts/sec	data	$\frac{(3.40)^m}{m!}e^{-3.40}$
0	$106/3600 = .029$	.033
1	$414/3600 = .115$	.113
2	$714/3600 = .198$	.193
3	$767/3600 = .213$	.219
4	$674/3600 = .187$	.186
5	$455/3600 = .126$	.126
6	$268/3600 = .074$	.072
7	$131/3600 = .036$	.035
8	$46/3600 = .013$	.015
9	$16/3600 = .0044$	.0056
10	$6/3600 = .0017$	.0019
11	$2/3600 = .0006$	.0006
12	$1/3600 = .0003$	.0002