Poisson Probability Distribution

Animal Health Risk Analysis
Canadian Food Inspection Agency
3851 Fallowfield Road
Ottawa, Ontario, Canada K2H 8P9

Analyse des risques de la santé des animaux
Agence canadienne d'inspection des aliments
3851, chemin Fallowfield
Ottawa (Ontario) Canada K2H 8P9

Canadian Food Inspection Agency (CFIA)

Agence canadienne d'inspection des aliments (ACIA)

Properties of the Poisson Distribution

The Poisson distribution describes the probability of the number of times that a random event will occur in a time or space interval under the conditions that

the probability of the event occurring is very small,

but the number of trials is very large so that the event actually occurs a few times.

The rate that the random event occurs is denoted by λ.

In order for the Poisson distribution to be an appropriate model for counting discrete points occuring in some sort of continuum, the following two assumptions must hold:

1) the number of events occurring in one part of the continuum should be statistically independent of the number of events occurring in another part of the continuum. In counting the number of red blood cells in diluted blood, the number of red blood cells in one square could be considered as statistically independent of the number of cells in another square.

and 2) the expected number of counts in a given part of the continuum should approach zero as its size approaches zero. For example, in observing blood cells one does not expect to find any in a very small area of a diluted specimen of blood.

Properties of the Poisson Distribution

When an event repeats itself after a time, for example, laboratory submissions of specimens of suspect hog cholera outbreaks, in a time interval of *t* time units, the event may occur zero, one, two, three, etc., times. Any process which satisfies the following assumptions is called a <u>Poisson process</u>.

1) the probability of the event occurring a specified number of times, say *x*, in a time interval of *t* time units does not change with time, that is, the probability of occurrence of the event is constant for any two intervals of time or space.

2) the probability of the event occurring *x* times in a time interval of *t* time units is unaffected by information concerning occurrences before the time interval began, that is the occurrence of the event in any interval is independent of its occurrence in any other interval.

3) if *h* > 0 is small, then the probability of the event occurring more than once in a time interval of *h* time units is negligible.

$$P(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

Properties of the Poisson Distribution

The Poisson distribution applies when an experiment yields an infinite number of possible counts in a continuous time interval or region of space where only the average number of counts per unit time or space is known. It often refers to rare events. A unit of time may be a minute, an hour, a day, a week while a unit of space may be a length, a distance, an area, a volume. For instance, the number of telephone calls per minute at a switchboard, the number of airplane arrivals per hour at an airport, the number of defects in a piece of furniture, the number of customer arrivals per half-hour, the number of defective electrical connections per mile of wiring in a city's power system are all Poisson random variables.

Properties of the Poisson Distribution

Examples in which discrete events occur in space or time:

The number of *Escherichia coli* bacteria in a sample of water. If a small quantity of the water is examined after thorough mixing of the water , then the probability of *x* of the bacteria being present in the sample is a Poisson probability.

The number of purebred Holstein cattle destined for export infected with Johne's disease. Nationally the disease occurs at the rate of 1 per 100,000 dairy cattle.

The number of cattle trucks arriving at a disease control inspection station in a country that has eradicated FMD but enforces vaccination. The rate is $\lambda t = 30$ cattle trucks per day from Monday to Friday, inclusive.

When counting numbers of red blood cells in diluted blood that occur in a specified rectangular area marked off in the field of view of a hemocytometer.

Another example would be the counting of the number of radioactive emissions from a radioactive source. Given that the period of observation is small in comparison with the radioactivity half life of the particle, the counts could be modeled by a Poission distribution.

Properties of the Poisson Distribution

Probability mass function:

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0,1,2,3,\ldots$$

Cumulative Distribution function:

$$F(x) = e^{-\lambda} \sum_{i=0}^{x} \frac{\lambda^i}{i!}$$

Parameter: λ

Domain: {0, 1, 2, 3, ...}

Mean (μ): λ (mean = variance, that is $\mu = \sigma^2$)

Variance (σ^2): λ

Mode: λ and λ-1 if λ is an integer, $\lfloor \lambda \rfloor$ otherwise

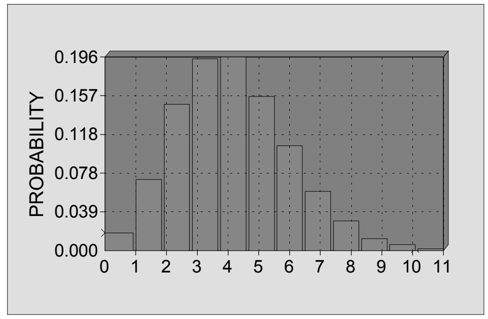
Coefficient of skewness (α_3): $1/\sqrt{\lambda}$

Coefficient of kurtosis (α_4): $3+1/\lambda$

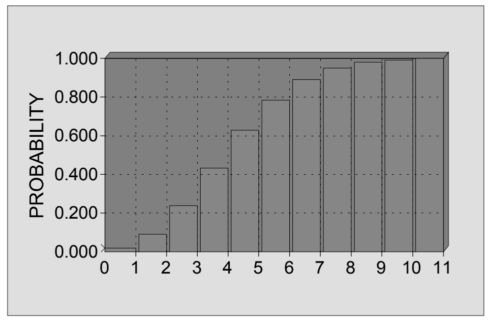
$\lfloor \rfloor$ = Greatest integer function

Properties of the Poisson Distribution

Probability mass function of POI(λ = 4) as an output of @RISK simulation



Cumulative distribution function of POI(λ = 4) as an output of @RISK simulation

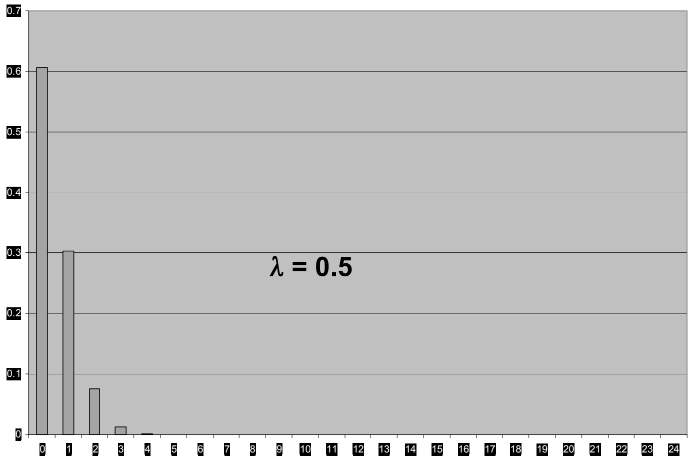


Properties of the Poisson Distribution

Bar graphs of the Poisson probabilities are given in the next four slides for λ = 0.5, 2, 5, and 10. As the mean (equal to the variance) increases, the distribution moves to the right and becomes more spread out and more symmetrical. In the first graph where λ = 0.5, the Microsoft Excel function for the Poisson distribution requires three inputs, x = the number of events, 0, 1, 2, etc., λ = 0.5 and a logical statement (true, false) as to whether the cumulative distribution function or the probability mass function is desired. Indicating “false” or “0", requests the probability mass function result for those values of x and λ.

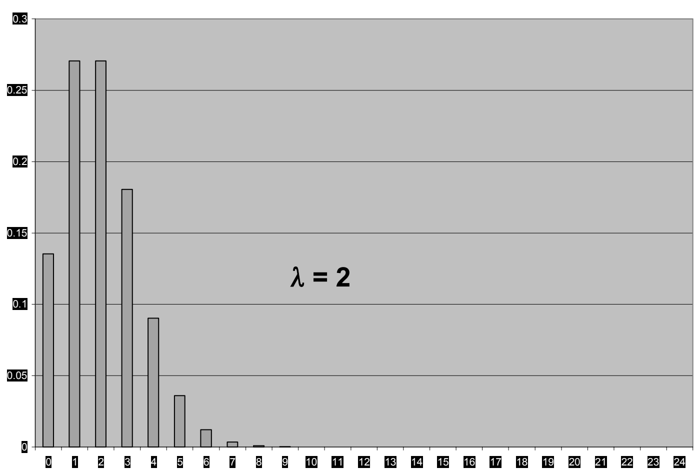
x	0.5
0	=POISSON(A11,B10,FALSE)
1	=POISSON(A12,B10,FALSE)
2	=POISSON(A13,B10,FALSE)
3	=POISSON(A14,B10,FALSE)
4	=POISSON(A15,B10,FALSE)
5	=POISSON(A16,B10,FALSE)
6	=POISSON(A17,B10,FALSE)

Properties of the Poisson Distribution



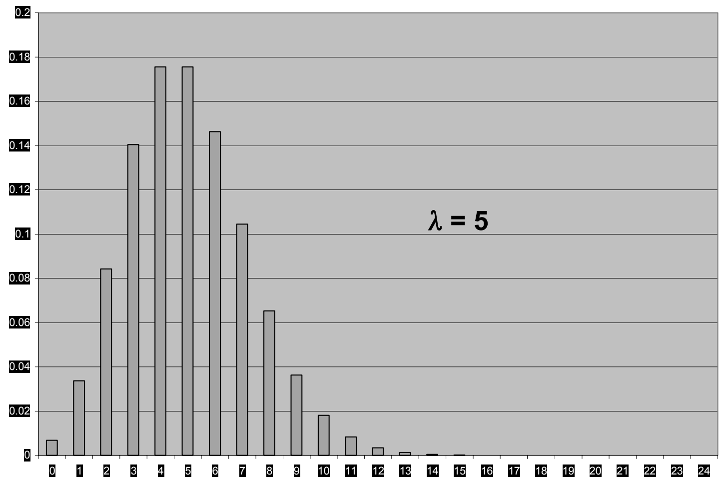
	Lamda Value	
x	0.5	2.0
0	0.60653066	0.135335
1	0.30326533	0.270671
2	0.075816332	0.270671
3	0.012636055	0.180447
4	0.001579507	0.090224
5	0.000157951	0.036089
6	1.31626E-05	0.01203
7	9.40183E-07	0.003437
8	5.87614E-08	0.000859
9	3.26452E-09	0.000191
10	1.63226E-10	3.82E-05
11	7.41937E-12	6.94E-06
12	3.0914E-13	1.16E-06
13	1.189E-14	1.78E-07
14	4.24643E-16	2.54E-08
15	1.41548E-17	3.39E-09
16	4.42337E-19	4.24E-10
17	1.30099E-20	4.99E-11
18	3.61386E-22	5.54E-12
19	9.51017E-24	5.83E-13
20	2.37754E-25	5.83E-14
21	5.66081E-27	5.56E-15
22	1.28655E-28	5.05E-16
23	2.79685E-30	4.39E-17

Properties of the Poisson Distribution



	Lamda Value	
x	0.5	2.0
0	0.60653066	0.135335
1	0.30326533	0.270671
2	0.075816332	0.270671
3	0.012636055	0.180447
4	0.001579507	0.090224
5	0.000157951	0.036089
6	1.31626E-05	0.01203
7	9.40183E-07	0.003437
8	5.87614E-08	0.000859
9	3.26452E-09	0.000191
10	1.63226E-10	3.82E-05
11	7.41937E-12	6.94E-06
12	3.0914E-13	1.16E-06
13	1.189E-14	1.78E-07
14	4.24643E-16	2.54E-08
15	1.41548E-17	3.39E-09
16	4.42337E-19	4.24E-10
17	1.30099E-20	4.99E-11
18	3.61386E-22	5.54E-12
19	9.51017E-24	5.83E-13
20	2.37754E-25	5.83E-14
21	5.66081E-27	5.56E-15
22	1.28655E-28	5.05E-16
23	2.79685E-30	4.39E-17

Properties of the Poisson Distribution



	Lamda Value	
x	**5.0**	**10.0**
0	0.006738	4.54E-05
1	0.03369	0.000454
2	0.084224	0.00227
3	0.140374	0.007567
4	0.175467	0.018917
5	0.175467	0.037833
6	0.146223	0.063055
7	0.104445	0.090079
8	0.065278	0.112599
9	0.036266	0.12511
10	0.018133	0.12511
11	0.008242	0.113736
12	0.003434	0.09478
13	0.001321	0.072908
14	0.000472	0.052077
15	0.000157	0.034718
16	4.91E-05	0.021699
17	1.45E-05	0.012764
18	4.01E-06	0.007091
19	1.06E-06	0.003732
20	2.64E-07	0.001866
21	6.29E-08	0.000889

Properties of the Poisson Distribution

λ = 10

	Lamda Value	
x	**5.0**	**10.0**
0	0.006738	4.54E-05
1	0.03369	0.000454
2	0.084224	0.00227
3	0.140374	0.007567
4	0.175467	0.018917
5	0.175467	0.037833
6	0.146223	0.063055
7	0.104445	0.090079
8	0.065278	0.112599
9	0.036266	0.12511
10	0.018133	0.12511
11	0.008242	0.113736
12	0.003434	0.09478
13	0.001321	0.072908
14	0.000472	0.052077
15	0.000157	0.034718
16	4.91E-05	0.021699
17	1.45E-05	0.012764
18	4.01E-06	0.007091
19	1.06E-06	0.003732
20	2.64E-07	0.001866
21	6.29E-08	0.000889
22	1.43E-08	0.000404
23	3.11E-09	0.000176

Horse-kick Example

A famous example of the Poisson distribution is data by von Bortkiewicz (1898) showing the chance of a cavalryman of the Prussian army being killed by a horse-kick in the course of a year. The data are from recordings of ten corps over a period of 20 years supplying 200 readings or 200 intervals.

x = Number of Deaths	Observed Number of Corps-Years in which x Fatalities Occurred
0	109
1	65
2	22
3	3
4	1
	200

There were 122 fatalities [109(0) + 65(1) + 22(2) + 3(3) + 1(4) with an observed fatality rate of 122/200 or 0.61 fatalities per corps-year. If each fixed interval is divided into n sub-intervals where n is very large and we assume that the probability of an occurrence or non-occurrence in a sub-interval is close to 1, then the probability of more that one occurrence is very small. The occurrence of an event in any sub-interval is independent of the occurrence in any other sub-interval. Thus this represents a <u>Poisson Process</u> in the fixed interval.

Horse-kick Example

The parameter λ is the expectation of the single Poisson random variable. Here, the parameter is λt, the expectation of a single observation by the average number of deaths per interval, 0.61. Hence, the Poisson distribution is:

$$P(X=x)=\frac{(\lambda t)^x e^{-\lambda t}}{x!}=\frac{(0.61)^x e^{-0.61}}{x!}$$

$$P(X=0)=\frac{(0.61)^0 e^{-0.61}}{0!}=0.5434$$

$$P(X=1)=\frac{(0.61)^1 e^{-0.61}}{1!}=0.3314$$

The Poisson-distribution prediction of horse-kick fatalities for the 200 corps-years, is therefore 108.7 corps-years with zero fatalities and 66.3 years with one fatality.

Dog Bite Example

According to the National Animal Health Authorities, the average number of humans bitten by dogs that necessitates quarantine of the dog for rabies is 2.0 per 100,000 population. In an area having a population of 200,000, what is the probability that there will be 0, between 4 and 8, and less than 3 dog bites per year that require rabies quarantine?

Letting X represent the number of dog bites requiring quarantine, one can assume that X is Poisson distributed especially as the rate is a rare event. For the area of interest, λ = *np* = (200,000)(0.00003) = 4.

The probability of no dog bites requiring quarantine is represented as P(X=0). The probability of between 4 and 8 dog bites requiring quarantine is $P(4 < POI(\lambda=4) < 8)$ while the probabilty of less than 4 such events is $P(X \leq 3)$.

$$P(X=0) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{4^0 e^{-4}}{0!} = 0.0183$$

$$P(4 < POI(\lambda) < 8) = \sum_{x=5}^{7} \frac{\lambda^x e^{-\lambda}}{x!} = \frac{4^5 e^{-4}}{5!} + \frac{4^6 e^{-4}}{6!} + \frac{4^7 e^{-4}}{7!} = 0.1563 + 0.1042 + 0.0595 = 0.3200$$

Dog Bite Example

From the Poisson distribution table of cumulative probabilities on the following slide, the probability $P(4 < POI(\lambda=4) < 8)$ can be obtained as F(X=7) - F(X=4) = 0.949 - 0.629 = 0.320.

To use the Cumulative Poisson Probability tables,

1) Find the λ = E(X) value column

2) For that column, the row representing the value of x gives the $P(X \leq x)$

3) List the values of X in the event of interest

4) Write the event as either the difference of two probabilities $P(X \leq a)$ - $P(X \leq b)$ or as 1 - $P(X \leq a)$.

From the Poisson distribution table of cumulative probabilities on the following slide, the probability $P(X \leq 3)$ can be obtained as F(X=3) , where λ = 4 and x = 3.

$$P(X \leq 3) = \sum_{x=0}^{3} \frac{\lambda^x e^{-\lambda}}{x!} = p(0) + p(1) + p(2) + p(3) = 0.0183 + 0.0733 + 0.1465 + 0.1954 = 0.4335$$

The Poisson Distribution $P(X \le x)$ $\lambda = E(X)$

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	0.905	0.819	0.741	0.670	0.607	0.549	0.497	0.449	0.407	0.368
1	0.995	0.982	0.963	0.938	0.910	0.878	0.844	0.809	0.772	0.736
2	1.000	0.999	0.996	0.992	0.986	0.977	0.966	0.953	0.937	0.920
3	1.000	1.000	1.000	0.999	0.998	0.997	0.994	0.991	0.987	0.981
4	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.996
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
0	0.333	0.301	0.273	0.247	0.223	0.202	0.183	0.165	0.150	0.135
1	0.699	0.663	0.627	0.592	0.558	0.525	0.493	0.463	0.434	0.406
2	0.900	0.879	0.857	0.833	0.809	0.783	0.757	0.731	0.704	0.677
3	0.974	0.966	0.957	0.946	0.934	0.921	0.907	0.891	0.875	0.857
4	0.995	0.992	0.989	0.986	0.981	0.976	0.970	0.964	0.956	0.947
5	0.999	0.998	0.998	0.997	0.996	0.994	0.992	0.990	0.987	0.983
6	1.000	1.000	1.000	0.999	0.999	0.999	0.998	0.997	0.997	0.995
7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
x	2.2	2.4	2.6	2.8	3	3.2	3.4	3.6	3.8	4
0	0.111	0.091	0.074	0.061	0.050	0.041	0.033	0.027	0.022	0.018
1	0.355	0.308	0.267	0.231	0.199	0.171	0.147	0.126	0.107	0.092
2	0.623	0.570	0.518	0.469	0.423	0.380	0.340	0.303	0.269	0.238
3	0.819	0.779	0.736	0.692	0.647	0.603	0.558	0.515	0.473	0.433
4	0.928	0.904	0.877	0.848	0.815	0.781	0.744	0.706	0.668	0.629
5	0.975	0.964	0.951	0.935	0.916	0.895	0.871	0.844	0.816	0.785
6	0.993	0.988	0.983	0.976	0.966	0.955	0.942	0.927	0.909	0.889
7	0.998	0.997	0.995	0.992	0.988	0.983	0.977	0.969	0.960	0.949
8	1.000	0.999	0.999	0.998	0.996	0.994	0.992	0.988	0.984	0.979
9	1.000	1.000	1.000	0.999	0.999	0.998	0.997	0.996	0.994	0.992
10	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.997
11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999
12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Abcess Example

If the probabillity of an abscess reaction from killed foot and mouth vaccine is 0.001, what is the probability of that in a herd of 2000 cattle there will be more than one animal with abscessation?

Let X represent the number of cattle with vaccine induced abscessation. Although X is Bernoulli distributed, X can be assumed to be Poisson distributed since the abscess reaction is a rare event.

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad where\ \lambda = np = 2000 \times 0.001 = 2$$

$$P(X > 1) = 1 - [P(X = 0) + P(X = 1)$$

$$1 - \left[\frac{2^0 e^{-2}}{0!} + \frac{2^1 e^{-2}}{1!} \right] = 1 - 3e^{-2} = 0.594$$

Bayesian Update Example

A risk assessment was conducted on the likelihood of foot and mouth disease outbreaks associated with the exposure of free-living swine to food waste in landfill sites. The estimated probabilities for nine frequencies of foot and mouth outbreaks per year were elaborated with the risk assessment (see table below). These represent the prior probabilities. The numerator of the Bayes' Rule answers the question, if the frequency of occurrence is correct, how likely is the evidence? The evidence (E) was represented as zero (0) occurrences of foot and mouth disease over T = 100 years of observation. $P(E|A_i)$ was estimated using the Poisson distribution expression in which *k* = 0 and T = 100 years.

Discrete form of Bayes' Rule:

$$P(A_i|E) = \frac{P(A_i)P(E|A_i)}{\sum_{i=1}^{9} P(A_i)P(E|A_i)},$$

$$E = \{k = 0 \ \ occurrences, T = 100\, years\}$$

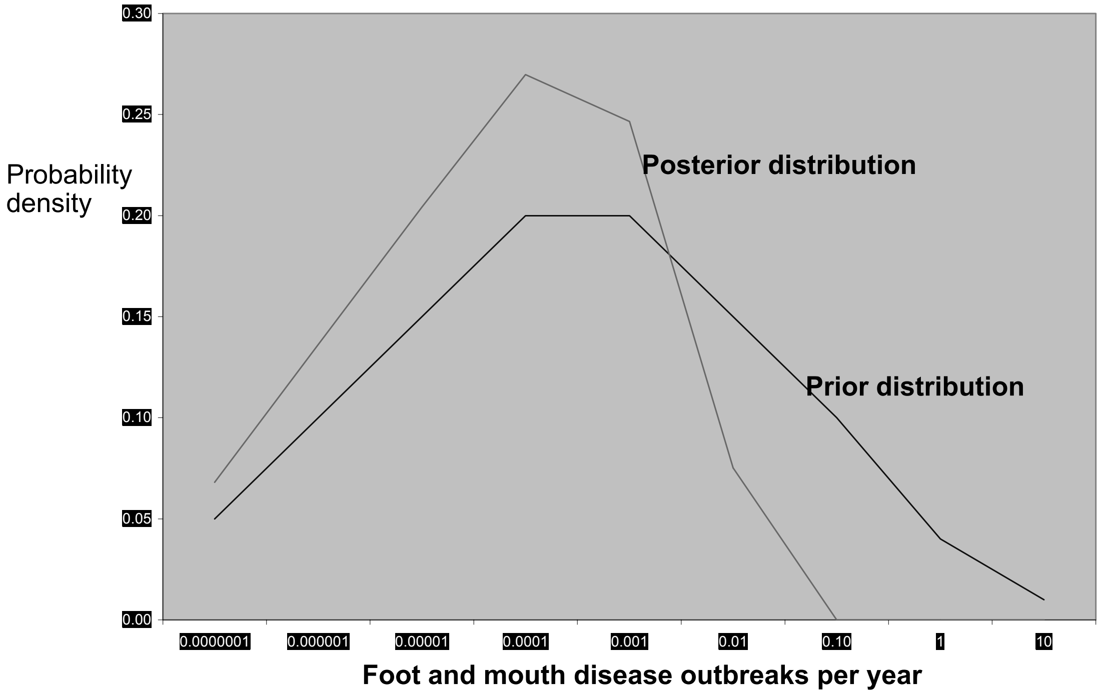
$$P(E|A_i) = \frac{(A_i T)^k}{k!} e^{-A_i T}.$$

Bayesian Update Example

Frequency (A_i)	Prior $P(A_i)$	$P(E\|A_i)$	$P(A_i)P(E\|A_i)$	Posterior $P(A_i\|E)$
0.0000001	0.05	1.0000	0.050	0.068
0.000001	0.10	0.9999	0.100	0.136
0.00001	0.15	0.9990	0.150	0.204
0.0001	0.20	0.9900	0.198	0.270
0.001	0.20	0.9048	0.181	0.247
0.01	0.15	0.3679	0.055	0.075
0.10	0.10	4.54E-05	4.54E-06	0.000
1	0.04	3.72E-44	1.49E-45	0.000
10	0.01	0	0	0.000
	1.00		0.734	1.000

Probability density curves are presented on the next slide to illustrate both prior and posterior distribution curves and the shifting to the left of the tail of the posterior distribution based on the evidence.

Bayesian Update Example



Relationship Between the Binomial and Poisson Distributions

In the binomial distribution, if *n* is large while the probability *p* of occurrence of an event is close to zero, so that *q* = 1-*p* is close to 1, the event is called a rare event. In pracitce an event is considered rare if the number of trials is at least 50 ($n \geq 50$) while *np* is less than 5. For such cases the binomial distribution is very closely approximated by the Poisson distribution with $\lambda = np$.

Relationship Between the Poisson and Normal Distributions

Since there is a relation between the binomial and normal distributions and between the binomial and Poisson distributions, there is in fact a relation between the Poisson and normal distributions. The Poisson distribuion approaches the normal distribuion as $\lambda \rightarrow \infty$.