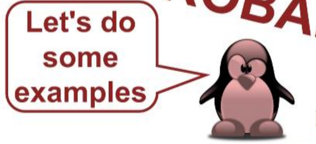
SE

POISSON PROBABILITY

STEP-BY-STEP EXAMPLES



$$p(x) = \frac{e^{-\bar{x}}\bar{x}^x}{x!}$$

Review of Poisson probability equations

When a process is Poisson distributed (i.e., random) and we know the mean number per unit time or area, these equations give the probabilities of seeing x or $x+1$ successes in that unit time or area.

$$p(x) = \frac{\mu^x e^{-\mu}}{x!} = \frac{\bar{x}^x e^{-\bar{x}}}{x!}$$

$$p(x+1) = \frac{\mu}{x+1}p(x) = \frac{\bar{x}}{x+1}p(x)$$

Example #1 - fish in nets

Let's consider a Example in which we cast a net into an ocean and we're estimating how many individuals of a rare fish species we expect to catch.

Each net potentially catches thousands of fish, but we know from experience that the average (i.e., mean) is 3.

What are the probabilities of nets with 0, 1, 2, 3, etc. fish?

Example #1 - fish in nets

Mean = 3.
What are the probabilities of nets with 0, 1, 2, 3, etc. fish?

$$p(x) = \frac{\mu^x e^{-\mu}}{x!} = \frac{3^x e^{-3}}{x!}$$

$$p(0) = \frac{3^x e^{-3}}{x!} = \frac{3^0 e^{-3}}{0!} = \frac{(1)(0.049787)}{(1)} = 0.049787$$

$$p(1) = \frac{3^x e^{-3}}{x!} = \frac{3^1 e^{-3}}{1!} = \frac{(3)(0.049787)}{(1)} = 0.149361$$

Example #1 - fish in nets

Mean = 3.
What are the probabilities of nets with 0, 1, 2, 3, etc. fish?

$$p(x) = \frac{\mu^x e^{-\mu}}{x!} = \frac{3^x e^{-3}}{x!}$$

$$p(2) = \frac{3^x e^{-3}}{x!} = \frac{3^2 e^{-3}}{2!} = \frac{(9)(0.049787)}{(2\times 1)} = \frac{(0.44808)}{(2)} = 0.2240$$

$$p(3) = \frac{3^x e^{-3}}{x!} = \frac{3^3 e^{-3}}{3!} = \frac{(27)(0.049787)}{(3\times 2\times 1)} = \frac{(1.344249)}{(6)} = 0.2240$$

Example #2 - amputees in town

Let's consider small towns in the US (and their medical needs)

How many amputees would we expect in a town of 5,000?

Each town *could* have hundreds.

The estimated frequency of major amputees in the US is 1 per 445 (*Ziegler-et al, 2008*), which would be 11.25 per 5,000.

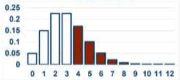
Example #1 - fish in nets

Cumulative probabilities of nets with 0, 1, 2, 3, etc. fish

p(0)=0.04979	p(x<1)=0.04979
p(1)=0.14936	p(x<2)=0.19915
p(2)=0.22404	p(x<3)=0.42319
p(3)=0.22404	p(x<4)=**0.64723**
p(4)=0.16803	p(x<5)=0.81526
p(5)=0.10082	p(x<6)=0.91608
p(6)=0.05041	p(x<7)=0.96649
p(7)=0.02160	p(x<8)=0.98810

What is the probability of a net with more than 3 fish?

p(x>3) =1-p(x<4)
=1-**0.64723**
=0.35277

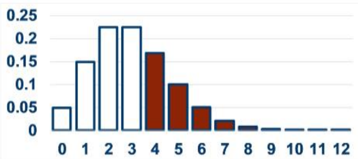


Example #1 - fish in nets

What is the probability of a net with more than 3 fish?

Can't sum all (never ends)

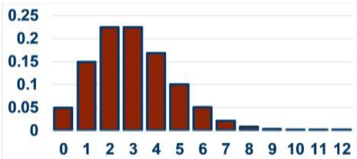
Can subtract p(0)-p(3) from 1



Example #1 - fish in nets

What are the probabilities of nets with 0, 1, 2, 3, etc. fish?

p(0)=0.049787
p(1)=0.149361
p(2)=0.224042
p(3)=0.224042
p(4)=0.168031
p(5)=0.100819
p(6)=0.050409
p(7)=0.021604



Example #1 - fish in nets

Mean = 3.
What are the probabilities of nets with 0, 1, 2, 3, etc. fish?

$p(2) = 0.2240$ $\qquad$ $p(3) = 0.2240$

$$p(x+1) = \frac{\mu}{x+1}p(x) \qquad x = 2,\ \mu = 3$$

$$p(3) = p(2) = \frac{3}{3}p(2)$$

Example #2 - amputees in town

Mean = 11.25.
What are the probabilities of towns with 5, 10, 15, 20, etc. amputees?

$$p(x) = \frac{\mu^x e^{-\mu}}{x!} = \frac{11.25^x e^{-11.25}}{x!}$$

$$p(5) = \frac{(11.25)^5 e^{-11.25}}{5!} = \frac{(180203.247)(1.3007297\times 10^{-5})}{(120)}$$

$$= \frac{(2.343957)}{(120)} = 0.0195$$

Example #2 - amputees in town

Mean = 11.25.
What are the probabilities of towns with 5, 10, 15, 20, etc. amputees?

$$p(x) = \frac{\mu^x e^{-\mu}}{x!} = \frac{11.25^x e^{-11.25}}{x!}$$

$$p(10) = \frac{(11.25)^{10} e^{-11.25}}{10!} = \frac{(32473210255)(1.3007297\times 10^{-5})}{(120)}$$

$$= \frac{(32473210255)(1.3007297\times 10^{-5})}{(3628800)} = 0.1164$$

Example #2 - amputees in town

Mean = 11.25.
What are the probabilities of towns with 5, 10, 15, 20, etc. amputees?

$$p(x) = \frac{\mu^x e^{-\mu}}{x!} = \frac{11.25^x e^{-11.25}}{x!}$$

$$p(15) = \frac{(11.25)^{15} e^{-11.25}}{15!}$$

5.85177x10[15]
1.30767x10[12]

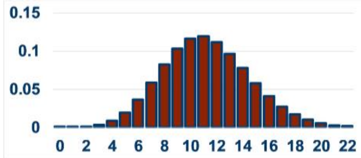
$$p(15) = \frac{(11.25)^5}{15\times 14\times 13\times 12\times 11}p(10) = 0.0582$$

Example #2 - amputees in town

What are the probabilities of towns with 5, 10, 15, 20, etc. amputees?

p(0)=0.00001
p(5)=0.0195
p(10)=0.1164
p(15)=0.0582
p(20)=0.0056

p(>15)?

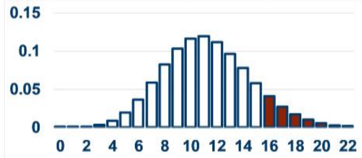


Example #2 - amputees in town

What's the probability of a town with more than 15 amputees? p(>15)?

Can't sum all (never ends)

Can subtract p(0)-p(15) from 1



Example #4 - is a pattern random?

Q: what if we had a table of event observation data from a scenario that someone claims is due to a random process?

A1: we can compare the observed values to the ones predicted from a Poisson distribution.

A2: we can compare the mean and variance of the observed values.

# events in period	# times observed
0	24
1	30
2	26
3	15
4	3
5	1
6+	0

Example #3 - estimating the mean from partial data

X=0
$$\frac{p(1)(1)}{p(0)} = \frac{28}{15}(1) = 1.8666 = \bar{x}$$

X=1
$$\frac{p(2)(2)}{p(1)} = \frac{27}{28}(2) = 1.9286 = \bar{x}$$

X=2
$$\frac{p(3)(3)}{p(2)} = \frac{17}{27}(3) = 1.8888 = \bar{x}$$

X=3
$$\frac{p(4)(4)}{p(3)} = \frac{8}{17}(4) = 1.8824 = \bar{x}$$

# events in period	# times observed
0	15
1	28
2	27
3	17
4	8
5+	4

$$\frac{p(x+1)(x+1)}{p(x)} = \bar{x}$$

Example #3 - estimating the mean from partial data

Use the shortcut to figure out the mean.

$$p(x+1) = \frac{\bar{x}}{x+1}p(x)$$

$$\frac{p(x+1)(x+1)}{p(x)} = \bar{x}$$

Keeping in mind that there is sampling and rounding error. Best practice would be to calculate for all steps (and favor most common).

# events in period	# times observed
0	15
1	28
2	27
3	17
4	8
5+	4

Example #3 - estimating the mean from partial data

What if we had a table of event observation data from a scenario that is due to a random process?

How could we figure out the mean?

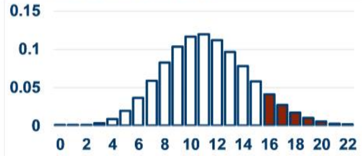
We can't just calculate a weighted average because the values in the "5+" category could be 5, 6,7, 8, etc. and we don't know which.

# events in period	# times observed
0	15
1	28
2	27
3	17
4	8
5+	4

Example #2 - amputees in town

0	0.000013	0.000013
1	0.000146	0.000159
2	0.000823	0.000982
3	0.003087	0.004069
4	0.008681	0.012750
5	0.019533	0.032283
6	0.036624	0.068908
7	0.058861	0.127768
8	0.082773	0.210541
9	0.103466	0.314007
10	0.116399	0.430406
11	0.119044	0.549450
12	0.111604	0.661054
13	0.096581	0.757635
14	0.077609	0.835244
15	0.058207	0.893451
16	0.040927	0.934378
17	0.027084	0.961462
18	0.016927	0.978389
19	0.010023	0.988412
20	0.005638	0.994050

What's the probability of a town with more than 15 amputees? p(>15)?
1-0.893451=0.106549, over 10%



Example #4 - is a pattern random?

A1: we can compare the **observed** values to the ones **predicted** from a Poisson distribution and 100 (20+40+21+14+4+1=100) observed periods.

# events in period	# times observed
0	**20**
1	**40**
2	**21**
3	**14**
4	**4**
5	**1**
6+	**0**

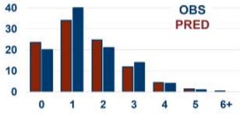
p(0)=0.2346 x 100 = **23.46**
p(1)=0.3401 x 100 = **34.01**
p(2)=0.2466 x 100 = **24.66**
p(3)=0.1192 x 100 = **11.92**
p(4)=0.0432 x 100 = **4.32**
p(5)=0.0125 x 100 = **1.25**
p(6+)=1-sum of previous=**0.38**

Mean of all these values is $\bar{x} = 1.45$.

0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,1,1,1,1,1,1,1,1,
1,1,1,1,1,1,1,1,1,1,1,1,1,1,
1,1,1,1,1,1,1,1,1,1,1,1,1,1,
1,1,1,1,2,2,2,2,2,2,2,2,2,2,
2,2,2,2,2,2,2,2,2,2,2,3,3,3,
3,3,3,3,3,3,3,3,3,3,3,4,4,4,
4,5

Example #4 - is a pattern random?

A1: we can compare the **observed** values to the ones **predicted** from a Poisson distribution and 100 observed periods.

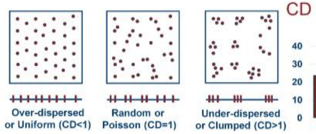


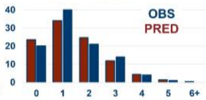
# events in period	# times observed	# times predicted	degree of mismatch
0	20	23.46	-3.46
1	40	34.01	+5.99
2	21	24.66	-3.66
3	14	11.92	+2.08
4	4	4.32	-0.32
5	1	1.25	-0.25
6+	0	0.38	-0.38

Example #4 - is a pattern random?

A2: we can compare the mean and variance of the observed values. From the data values, $\bar{x} = 1.450$ and $s^2 = 1.300$. The coefficient of dispersion is their ratio.

$$CD = \frac{s^2}{\bar{x}} = \frac{1.300}{1.450} = 0.897$$





Example #4 - is a pattern random?

Comparing observed values & predicted values to test claims of randomness.

- If they match, distribution *probably* Poisson/random.
- If they don't match, distribution not Poisson/random.

But ... there will always be some mismatch (sampling error, rounding).

How much mismatch is too much ... non-random?

Need a technique like the chi-squared analysis.

Definition equation $\qquad p(x) = \frac{\mu^x e^{-\mu}}{x!} = \frac{\bar{x}^x e^{-\bar{x}}}{x!}$

Shortcut equation $\qquad p(x+1) = \frac{\mu}{x+1}p(x) = \frac{\bar{x}}{x+1}p(x)$

See the companion video (linked in description and at the end) for more detail about these equations and more about applications of the Poisson probability distribution.

StatsExamples.com