The Poisson Distribution

The Poisson Probability Distribution

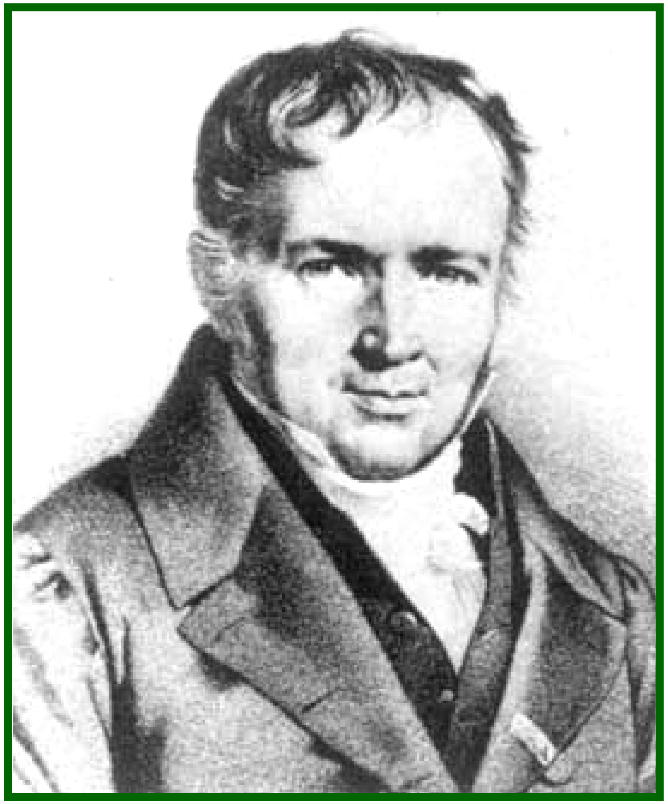
- "Researches on the probability of criminal civil verdicts" 1837
- Looked at the form of the binomial distribution

When the Number of Trials is Large.

- He derived the cumulative Poisson distribution as the

Limiting case of the Binomial When the Chance of Success Tends to Zero.

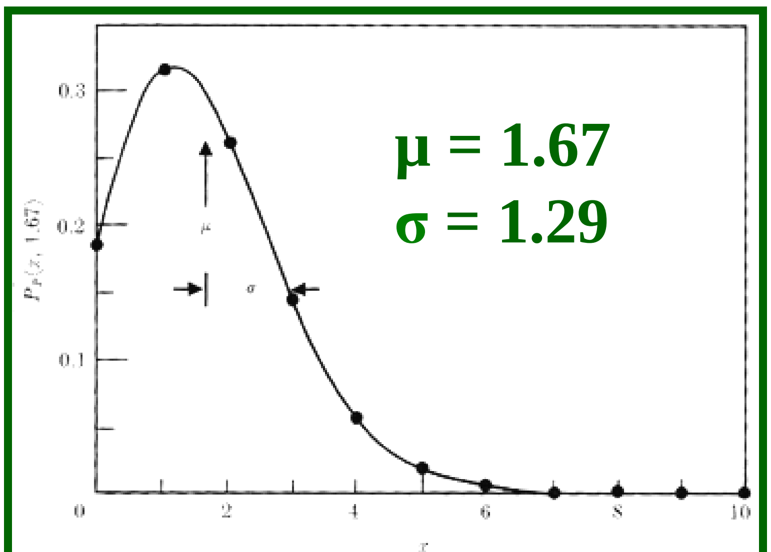
Simeon Denis Poisson

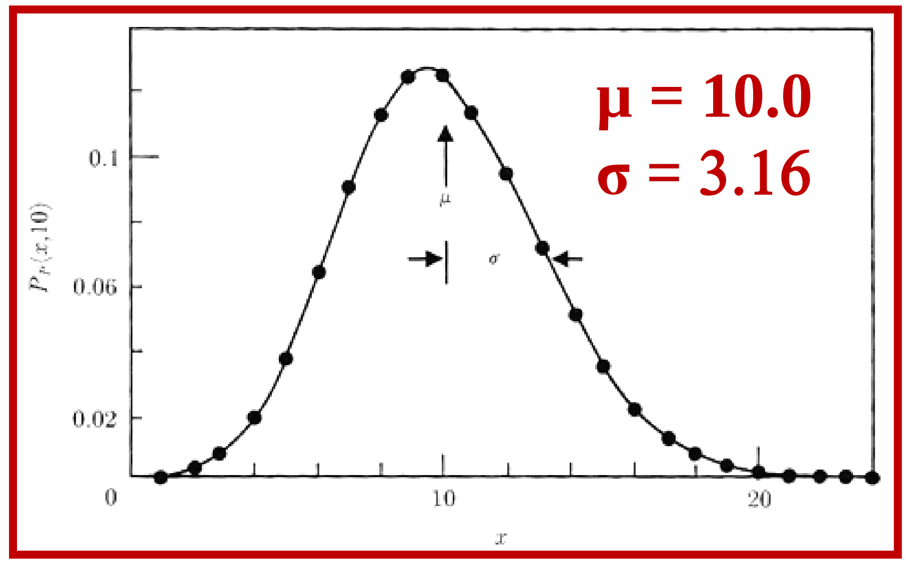


The Poisson Distribution

- **<u>Poisson Distribution</u>**: An approximation to the binomial distribution for ***<u>the SPECIAL CASE</u>*** ***when the average number (mean μ) of successes is very much smaller than the possible number n.*** *i.e.* $\mu \ll n$ because $p \ll 1$.
- This distribution is ***<u>NOT</u>*** necessarily symmetric!
- Data are usually bounded on one side & not the other.

An advantage of this distribution is that $\sigma^2 = \mu$





The Poisson Distribution ***Models Counts*.**

- If events happen at a constant rate over time, the **Poisson Distribution** gives

The Probability of X Number of Events Occurring in a time T.

- This distribution tells us the

Probability of All Possible Numbers of Counts, from 0 to Infinity.

- If **X= #** of counts per second, then the Poisson probability that **X = r** (a particular count) is:

$$\boxed{p(X = r) = \frac{\lambda^r e^{-\lambda}}{r!}}$$

λ ≡ the average number of counts per second.

Example 1

We introduced the Binomial distribution by considering the following scenario.

A worn machine is known to produce 10% defective components. If the random variable X *is the number of defective components produced in a run of 3* components, find the probabilities that X *takes the values 0 to 3.*

Suppose now that a similar machine which is known to produce 1% defective components is used for a production run of 40 components. We wish to calculate the probability that two defective items are produced. Essentially we are assuming that $X \sim B(40, 0.01)$ *and are asking for* $P(X = 2)$. *We use both* the Binomial distribution and its Poisson approximation for comparison.

Solution

Using the Binomial distribution we have the solution

$$\mathrm{P}(X=2) \;=\; C_{40}^{2}\cdot 0.99^{40-2}\cdot 0.01^{2} \;=\frac{40\cdot 39}{1\cdot 2}0.99^{38}\cdot 0.01^{2} = \;0.0532$$

Note that the arithmetic involved is unwieldy. Using the Poisson approximation we have the solution

$$\mathrm{P}(X=2) \;= e^{-0.4}\,\frac{0.4^{2}}{2!} = \;0.0532$$

Note that the arithmetic involved is simpler and the approximation is reasonable.

Example 2

A manufacturer produces light-bulbs that are packed into boxes of 100. If quality control studies indicate that 0.5% of the light-bulbs produced are defective, what percentage of the boxes will contain:

(a) no defective.

(b) 2 or more defectives?

Solution

As n is large and p, the P(defective bulb), is small, use the Poisson approximation to the Binomial probability distribution.

If X = number of defective bulbs in a box, then

$$X \sim P(\mu) \text{ where } \mu = n \times p = 100 \times 0.005 = 0.5$$

Hence,

(a) $P(X = 0) = \dfrac{e^{-0.5}(0.5)^0}{0!} = \dfrac{e^{-0.5}(1)}{1} = 0.6065 \approx 61\%$

(b) $P(X = 2 \text{ or more}) = P(X = 2) + P(X = 3) + P(X = 4) + \ldots$

But easier to consider,

$$P(X \geq 2) = 1 - [P(X = 0) + P(X = 1)]$$

$$P(X = 1) = \frac{e^{-0.5}(0.5)^1}{1!} = \frac{e^{-0.5}(0.5)}{1} = 0.3033$$

i.e. $\quad P(X \geq 2) = 1 - [0.6065 + 0.3033] = 0.0902 \approx 9\%$

Example 3

Suppose that it has been observed that, on average, 180 cars per hour pass a specified point on a particular road in the morning rush hour. Due to impending roadworks it is estimated that congestion will occur closer to the city centre if more than 5 cars pass the point in any one minute. What is the probability of congestion occurring?

Solution

Let X be the random variable X = number of cars arriving in any minute. We need to calculate the probability that more than 5 cars arrive in any one minute. Note that in order to do this we need to convert the information given on the average rate (cars arriving per hour) into a value for λ (cars arriving per minute). This gives the value $\lambda = 3$. Using $\lambda = 3$ to calculate the required probabilities gives:

r	0	1	2	3	4	5	Sum
$P(X = r)$	0.04979	0.149361	0.22404	0.22404	0.168031	0.10082	0.91608

To calculate the required probability we note that

$$P(\text{more than 5 cars arrive in one minute}) = 1 - P(\text{5 cars or less arrive in one minute})$$

Thus

$$\begin{aligned} P(X > 5) &= 1 - P(X \leq 5) \\ &= 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) - P(X = 4) - P(X = 5) \end{aligned}$$

Then $P(\text{more than } 5) = 1 - 0.91608 = 0.08392 = 0.0839$

Example 4

The mean number of bacteria per millilitre of a liquid is known to be 6. Find the probability that in 1 ml of the liquid, there will be:

(a) 0,

(b) 1,

(c) 2,

(d) 3,

(e) less than 4,

(f) 6 bacteria.

Solution

Here we have an *average rate of occurrences* but no estimate of the probability so it looks as though we have a Poisson distribution with $\lambda = 6$. Using the formula above we have:

(a) $P(X = 0) = e^{-6}\frac{6^0}{0!} = 0.00248$. That is the probability of having no bacteria in 1 ml of liquid is 0.00248

(b) $P(X = 1) = \frac{\lambda}{1} \times P(X = 0) = 6 \times 0.00248 = 0.0149$. That is the probability of having 1 bacteria in 1 ml of liquid is 0.0149

(c) $P(X = 2) = \frac{\lambda}{2} \times P(X = 1) = \frac{6}{2} \times 0.01487 = 0.0446$. That is the probability of having 2 bacteria in 1 ml of liquid is 0.0446

(d) $P(X = 3) = \frac{\lambda}{3} \times P(X = 2) = \frac{6}{3} \times 0.04462 = 0.0892$. That is the probability of having 3 bacteria in 1 ml of liquid is 0.0892

(e) $P(X < 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 0.1512$

(f) $P(X = 6) = e^{-6}\dfrac{6^6}{6!} = 0.1606$

Note that in working out the first 6 answers, which link together, all the figures were kept in the calculator to ensure accuracy. Answers were rounded off when written down. Never copy down answers correct to, say, 4 decimal places and then use those rounded figures to calculate the next figure as rounding-off errors will become greater at each stage. In the example above you would get answers 0.0025, 0.0150, 0.0450, 0.9000 and $P(X < 4) = 0.1525$. The difference is not great but could be very significant.

Example 5

A Council is considering whether to base a recovery vehicle on a stretch of road to help clear incidents as quickly as possible. The road concerned carries over 5000 vehicles during the peak rush hour period. Records show that, on average, the number of incidents during the morning rush hour is 5. The Council won’t base a vehicle on the road if the probability of having more than 5 incidents in any one morning is less than 30%. Based on this information should the Council provide a vehicle?

We need to calculate the probability that more than 5 incidents occur i.e. $P(X > 5)$. To find this we use the fact that $P(X > 5) = 1 - P(X \leq 5)$. Now, for this example:

$$P(X = r) = e^{-5}\frac{5^r}{r!}$$

Writing answers to 5 d.p. gives:

$$\begin{aligned}
P(X = 0) &= e^{-5}\frac{5^0}{0!} = 0.00674 \\
P(X = 1) &= 5 \times P(X = 0) = 0.03369 \\
P(X = 2) &= \frac{5}{2} \times P(X = 1) = 0.08422 \\
P(X = 3) &= \frac{5}{3} \times P(X = 2) = 0.14037 \\
P(X = 4) &= \frac{5}{4} \times P(X = 3) = 0.17547 \\
P(X = 5) &= \frac{5}{5} \times P(X = 4) = 0.17547
\end{aligned}$$

$$\begin{aligned}
P(X \leq 5) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) \\
&= 0.61596
\end{aligned}$$

The probability of more than 5 incidents is $P(X > 5) = 1 - P(X \leq 5) = 0.38403$. That is, the probability of having more than 5 incidents is 38.4% (to 3 s.f.) so the Council should provide a vehicle.

Mean and Variance for the Poisson Distribution

- It's easy to show that for this distribution,

The _Mean_ is: $\boxed{\mu = \lambda}$

- Also, it's easy to show that **The _Variance_ is:**

$$\boxed{\sigma^2 = \lambda}$$

$\Rightarrow$ **The _Standard Deviation_ is:** $\boxed{\sigma = \sqrt{\lambda}}$

$\Rightarrow$ For a Poisson Distribution, the variance and mean are equal!