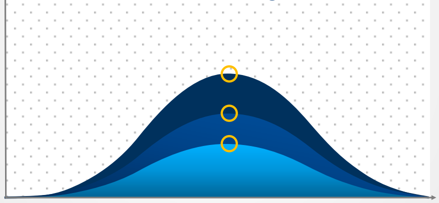
The Poisson Distribution

Md. Pear Hossain
Lecturer of Statistics
BSMRSTU, Bangladesh



Introduction

- In this lesson, we learn about another specially named discrete probability distribution, namely the **Poisson distribution**.

Objectives

- To learn the situation that makes a discrete random variable a Poisson random variable.
- To learn a heuristic derivation of the probability mass function of a Poisson random variable.
- To learn how to use the Poisson pmf to calculate probabilities for a Poisson random variable.
- To learn how to use a standard Poisson cumulative probability table to calculate probabilities for a Poisson random variable.
- To explore the key properties, such as the moment-generating function, mean and variance, of a Poisson random variable.
- To learn how to use the Poisson distribution to approximate binomial probabilities.
- To understand the steps involved in each of the proofs in the lesson.
- To be able to apply the methods learned in the lesson to new problems.

Situation

- Let the discrete random variable X denote the number of times an event occurs in an interval of time (or space). Then X *may* be a Poisson random variable with $x = 0, 1, 2, \ldots$

Examples

- Let X equal the number of typos on a printed page.
  (This is an example of an interval of space — the space being the printed page.)
- Let X equal the number of cars passing through the intersection of Dhaka-Khulna Highway and BSMRSTU road at Ghonapara in one minute.
  (This is an example of an interval of time — the time being one minute.)
- Let X equal the number of Hilsha caught in a squid driftnet.
  (This is again an example of an interval of space — the space being the squid driftnet.)
- Let X equal the number of customers at an ATM in 10-minute intervals.
- Let X equal the number of students arriving during office hours.

Definition.

If X is a Poisson random variable, then the probability mass function is:

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

for $x = 0, 1, 2, \ldots$ and $\lambda > 0$, where λ will be shown later to be both the mean and the variance of X.

"Derivation" of the pmf

Poisson distribution can be derived from the binomial distribution under the following conditions:

i. p, the probability of success in a Bernoulli trial is very small, that is $p \to 0$.

ii. n, the number of trials is very large, that is $n \to \infty$.

iii. $np = \lambda$ is finite constant, that is average number of success is finite.

Under this condition we have

$$p = \frac{\lambda}{n}$$

Hence,

$$q = 1 - p = 1 - \frac{\lambda}{n}$$

The probability function of binomial variate X with parameters n and p is

$$p(x) = \binom{n}{x} p^x q^{n-x}$$

$$p(x) = \frac{n!}{x!(n-x)!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$p(x) = \frac{n!}{x!(n-x)!}\lambda^x\left(\frac{1}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$p(x) = \frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}\frac{n!}{n^x(n-x)!} \qquad (1)$$

Now for fixed x;

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-x} = 1$$

and

$$\lim_{n\to\infty}\frac{n!}{n^x(n-x)!} = \lim_{n\to\infty}\frac{n(n-1)\ldots\{n-(x-1)\}(n-x)!}{n^x(n-x)!}$$

$$= \lim_{n\to\infty}\frac{n(n-1)\ldots\{n-(x-1)\}}{n^x}$$

$$= \lim_{n\to\infty}\frac{n\times n.\left(1-\frac{1}{n}\right)\times n\left(1-\frac{2}{n}\right)\times\cdots\times n\left\{1-\frac{(x-1)}{n}\right\}}{n^x}$$

$$= \lim_{n\to\infty}\frac{n^x\left(1-\frac{1}{n}\right)\times\left(1-\frac{2}{n}\right)\times\cdots\times\left\{1-\frac{(x-1)}{n}\right\}}{n^x}$$

$$= \lim_{n\to\infty}\left(1-\frac{1}{n}\right)\times\left(1-\frac{2}{n}\right)\times\cdots\times\left\{1-\frac{(x-1)}{n}\right\}$$

$= 1$

while,

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$$

Hence from (1) for large n we get

$$\lim_{n\to\infty} p(x) = \lim_{n\to\infty}\frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}\frac{n!}{n^x(n-x)!}$$

$$\lim_{n\to\infty} p(x) = \frac{\lambda^x}{x!}e^{-\lambda}\times 1\times 1 = \frac{e^{-\lambda}\lambda^x}{x!}$$

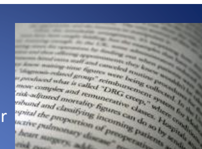
Which is the probability mass function of Poisson distribution.

Definition of exponential function

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n$$

$$e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}$$

Finding Poisson Probabilities



Let X equal the number of typos on a printed page with a mean of 3 typos per page. What is the probability that a randomly selected page has **at least one typo** on it?

- **Solution.** We can find the requested probability directly from the pmf The probability that X is at least one is:
- $P(X \geq 1) = 1 - P(X = 0)$
  Therefore, using the pmf to find $P(X = 0)$, we get:
- $P(X \geq 1) = 1 - \frac{e^{-3}3^0}{0!} = 1 - e^{-3} = 1 - 0.0498 = 0.9502$
  That is, there is just over a 95% chance of finding at least one typo on a randomly selected page when the average number of typos per page is 3.

What is the probability that a randomly selected page has **at most one typo** on it?

- **Solution.** The probability that X is at most one is:
- $P(X \leq 1) = P(X = 0) + P(X = 1)$
- Therefore, using the pmf, we get:
- $P(X \leq 1) = \frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} = e^{-3} + 3e^{-3} = 4e^{-3} = 4(0.0498) = 0.1992$
- That is, there is just under a 20% chance of finding at most one typo on a randomly selected page when the average number of typos per page is 3.

Cumulative Poisson probabilities using table

To find cumulative Poisson probabilities using table, do the following:

- Find the column headed by the relevant $\boldsymbol{\lambda}$.
- Find the $\boldsymbol{x}$ in the first column on the left for which you want to find $F(x) = P(X \leq x)$.

Table A.2 Cumulative Poisson Probabilities

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

		λ									
		.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
	0	.905	.819	.741	.670	.607	.549	.497	.449	.407	.368
	1	.995	.982	.963	.938	.910	.878	.844	.809	.772	.736
	2	1.000	.999	.996	.992	.986	.977	.966	.953	.937	.920
x	3		1.000	1.000	.999	.998	.997	.994	.991	.987	.981
	4				1.000	1.000	1.000	.999	.999	.998	.996

		λ										
		2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	15.0	20.0
	0	.135	.050	.018	.007	.002	.001	.000	.000	.000	.000	.000
	1	.406	.199	.092	.040	.017	.007	.003	.001	.000	.000	.000
	2	.677	.423	.238	.125	.062	.030	.014	.006	.003	.000	.000
	3	.857	.647	.433	.265	.151	.082	.042	.021	.010	.000	.000
	4	.947	.815	.629	.440	.285	.173	.100	.055	.029	.001	.000
	5	.983	.916	.785	.616	.446	.301	.191	.116	.067	.003	.000
	6	.995	.966	.889	.762	.606	.450	.313	.207	.130	.008	.000
	7	.999	.988	.949	.867	.744	.599	.453	.324	.220	.018	.001
	8	1.000	.996	.979	.932	.847	.729	.593	.456	.333	.037	.002
	9		.999	.992	.968	.916	.830	.717	.587	.458	.070	.005
	10		1.000	.997	.986	.957	.901	.816	.706	.583	.118	.011
	11			.999	.995	.980	.947	.888	.803	.697	.185	.021
	12			1.000	.998	.991	.973	.936	.876	.792	.268	.039
	13				.999	.996	.987	.966	.926	.864	.363	.066
	14				1.000	.999	.994	.983	.959	.917	.466	.105
	15					.999	.998	.992	.978	.951	.568	.157
	16					1.000	.999	.996	.989	.973	.664	.221
	17						1.000	.998	.995	.986	.749	.297
	18							.999	.998	.993	.819	.381
x	19							1.000	.999	.997	.875	.470
	20								1.000	.998	.917	.559
	21									.999	.947	.644
	22									1.000	.967	.721
	23										.981	.787

*Source JL Devore

Let's try it out on an example.

If X equals the number of typos on a printed page with a mean of 3 typos per page, what is the probability that a randomly selected page has **four typos** on it?

- **Solution.** The probability that a randomly selected page has four typos on it can be written as $P(X = 4)$. We can calculate $P(X = 4)$ by subtracting $P(X \le 3)$ from $P(X \le 4)$.
- To find $P(X \le 3)$ and $P(X \le 4)$ using the Poisson table, we:
  1. Find the column headed by $\lambda = 3$.
  2. Find the 3 in the first column on the left, since we want to find $F(3) = P(X \le 3)$.
  3. Find the 4 in the first column on the left, since we want to find $F(4) = P(X \le 4)$.
- Now, all we need to do is
  1. read the probability value where the $\boldsymbol{\lambda} = \mathbf{3}$ column and the $x = 3$ row intersect, and
  2. read the probability value where the $\boldsymbol{\lambda} = \mathbf{3}$ column and the $x = 4$ row intersect. What do you get?

- The cumulative Poisson probability table tells us that finding $P(X \le 4) = 0.815$ and $P(X \le 3) = 0.647$.
- Therefore:

  $P(X = 4) = P(X \le 4) - P(X \le 3) = 0.815 - 0.647 = 0.168$
- That is, there is about a 17% chance that a randomly selected page would have four typos on it.
- Since it wouldn't take a lot of work in this case, you might want to verify that you'd get the same answer using the Poisson pmf.

Properties of Poisson distribution

Theorem. The probability mass function:

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

for a Poisson random variable X is a valid pmf.

Proof.

a) $p(x) > 0$ because $\lambda^{x} > 0,\ e^{-\lambda} > 0,\ x! > 0$

b) $\sum_{x=0}^{\infty} p(x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^{x}}{x!} = e^{-\lambda}e^{\lambda} = 1$

MGF

Theorem. The moment generating function of a Poisson random variable X is:

$$M_X(t) = e^{\lambda(e^{t}-1)}\ for\ -\infty < t < \infty.$$

Proof.

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx}.\frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^{t})^{x}}{x!} = e^{-\lambda}e^{\lambda e^{t}} = e^{\lambda e^{t}-\lambda} = e^{\lambda(e^{t}-1)}$$

Mean

Theorem. The mean of a Poisson random variable X is λ.

Proof. *As we know that the $r-th$ moment about origin obtained from $M_{X(t)}$ is*

$$\left.\frac{d^{r}M_X(t)}{dt^{r}}\right|_{t=0}$$

Taking $r = 1$ we get

$$\frac{dM_X(t)}{dt} = \frac{d}{dt}e^{\lambda(e^{t}-1)} = e^{\lambda(e^{t}-1)}\frac{d}{dt}\lambda(e^{t}-1) = e^{\lambda(e^{t}-1)}\lambda e^{t}$$
$$= \lambda e^{\lambda(e^{t}-1)+t}$$

Hence the first raw moment of Poisson distribution is

$$\mu' = \left.\frac{dM_X(t)}{dt}\right|_{t=0} = \lambda e^{\lambda(e^{0}-1)+0} = \lambda e^{\lambda(1-1)} = \lambda$$

Therefore, Mean= $\mu' = \lambda$.

Variance

Theorem. The variance of a Poisson random variable X is λ.

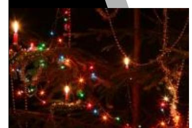
Proof. As we know that the $r - th$ *moment about origin obtained from* $M_{X(t)}$ *is*

$$\left.\frac{d^r M_X(t)}{dt^r}\right|_{t=0}$$

Taking $r = 2$ *we get*

……

Approximating the Binomial Distribution



Example. Five percent (5%) of beautification light bulbs manufactured by a company are defective. The company's Quality Control Manager is quite concerned and therefore randomly samples 100 bulbs coming off of the assembly line. Let *X* denote the number in the sample that are defective. What is the probability that the sample contains at most three defective bulbs?

- **Solution.** Can you convince yourself that *X* is a binomial random variable?