Binomial and Poisson Distribution

By Prateek Singla

&

Avtar Singh

What is a Probability Distribution?

- Listing of all possible outcomes of an experiment and probability of happening each of those events.
- Characteristics of a Probability distribution-
  - Probability of a particular outcome can be between 0 and 1, both inclusive.
  - The outcomes are mutually exclusive events.
  - The list is completely exhaustive. So, sum of all probabilities of various events would always be 1.

Binomial Distribution

- It is a discrete probability distribution.
- Binomial Probability is calculated by following general formula-

$$\boxed{P(X) = {}^{n}C_{x}\, p^{x}\, q^{(n-x)}}$$

Where, n = number of trials

x = number of success

p = Probability of success

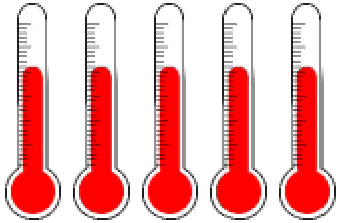
q = Probability of failure = 1 – p

Requirements for a Binomial Distribution

- ❖ Random Experiment must involve n identical trials.
- ❖ As the word “Binomial” suggests, each trail should have only 2 possible outcomes, denoted as “Success” or “Failure”.
- ❖ Each trial is independent of the previous trials.
- ❖ The probability of success denoted by p, does not change from trial to trial. The probability of failure is 1-p and it is also fixed from trial to trial.

A Simple Managerial Problem

ABS Laboratories is making a thermometer with 20% defect rate. If we select 5 randomly chosen thermometers at end of manufacturing process, what is the probability of having exactly 1 defective thermometer in our sample?



Hint – The answer is NOT 1/5 or 0.2 or 20 %

A Simple Managerial Problem

Here is a possible list of outcomes-

Trail	1	2	3	4	5
Outcome	F	F	S	F	F
Outcome	S	F	F	F	F
Outcome	F	F	F	S	F
Outcome	F	S	F	F	F
Outcome	F	F	F	F	S

A Simple Managerial Problem

Probability of success (Thermometer being defective) = 0.2
Probability of failure = 1 – 0.2 = 0.8

Trail	1	2	3	4	5	Probability
Outcome	0.8	0.8	0.2	0.8	0.8	0.08192
Outcome	0.2	0.8	0.8	0.8	0.8	0.08192
Outcome	0.8	0.8	0.8	0.2	0.8	0.08192
Outcome	0.8	0.2	0.8	0.8	0.8	0.08192
Outcome	0.8	0.8	0.8	0.8	0.2	0.08192
						0.4096

A Simple Managerial Problem

- It is given in problem that number of randomly selected samples is 5, therefore, $n = 5$
- It is also given that probability of a thermometer being defective is 20% or 0.2, therefore-

  $p = 0.2$, $q = 1 – p = 1 – 0.2 =0.8$
- We have to find probability of exactly one being defective, therefore $x = 1$, Using formula-

$P(X) = {}^{n}C_{x}\, p^{x}\, q^{(n-x)}$ » $P(1) = {}^{5}C_{1}\, (0.2)^{1}\, (0.8)^{(5-1)} = 0.4096$

Hence, probability of finding exactly 1 defective thermometer out of 5 is approximately 41%

Tabulating Probability Distribution

Calculation for all possible cases of success/failure.

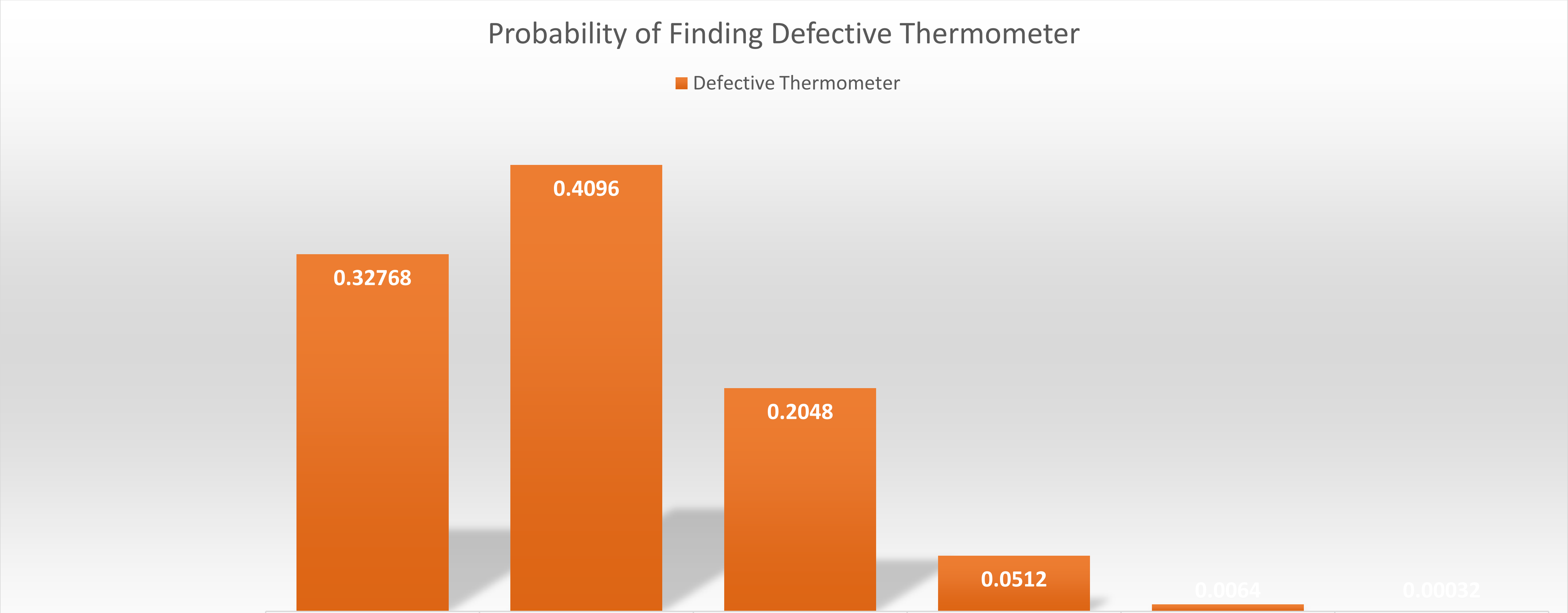
n	#successes	#failures	Calculation	Probability	Defective
5	0	5	${}^{5}C_{0} \text{ x } (0.2)^{0} \text{ x } (0.8)^{5}$	0.32768	0
5	1	4	${}^{5}C_{1} \text{ x } (0.2)^{1} \text{ x } (0.8)^{4}$	0.4096	1
5	2	3	${}^{5}C_{2} \text{ x } (0.2)^{2} \text{ x } (0.8)^{3}$	0.2048	2
5	3	2	${}^{5}C_{3} \text{ x } (0.2)^{3} \text{ x } (0.8)^{2}$	0.0512	3
5	4	1	${}^{5}C_{4} \text{ x } (0.2)^{4} \text{ x } (0.8)^{1}$	0.0064	4
5	5	0	${}^{5}C_{5} \text{ x } (0.2)^{5} \text{ x } (0.8)^{0}$	0.00032	5
				1	

Using Binomial Table

n = 5									
x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	0.590	0.328	0.168	0.078	0.031	0.010	0.002	0.000	0.000
1	0.328	0.410	0.360	0.259	0.156	0.077	0.028	0.006	0.000
2	0.073	0.205	0.309	0.346	0.313	0.230	0.132	0.051	0.008
3	0.008	0.051	0.132	0.230	0.313	0.346	0.309	0.205	0.073
4	0.000	0.006	0.028	0.077	0.156	0.259	0.360	0.410	0.328
5	0.000	0.000	0.002	0.010	0.031	0.078	0.168	0.328	0.590

Page 781 of Book

Graphing Probability Distribution



	Zero	One	Two	Three	Four	Five
Defective Thermometer	0.32768	0.4096	0.2048	0.0512	0.0064	0.00032

Mean of Binomial Distribution

n	p	q
5	0.2	0.8

Mean or expected value for Binomial distribution is given as-

$$\mu = np$$

Therefore, $\mu = 5 \times 0.2$

Mean = 1

Variance of Binomial Distribution

n	p	q
5	0.2	0.8

Variance of Binomial Distribution is given as-

$\sigma^2 = npq$

$= 5 \times 0.2 \times 0.8$

$= 0.8$

Standard Deviation of Binomial Distribution

n	p	q
5	0.2	0.8

Variance of Binomial Distribution is given as-

$$\sigma = \sqrt{npq}$$

$$= \sqrt{5 \times 0.2 \times 0.8}$$

$$= 0.8944$$

Final Results of Problem

<table>
<tr><th>n</th><th>p</th><th>q</th></tr>
<tr><td>5</td><td>0.2</td><td>0.8</td></tr>
<tr><td colspan="2">Probability of finding one defective thermometer</td><td>40.96 %</td></tr>
<tr><td>Mean</td><td>Variance</td><td>Standard Deviation</td></tr>
<tr><td>1</td><td>0.8</td><td>0.8944</td></tr>
</table>

Problem on Binomial Fitting

The following data shows the number of seeds germinating out of 12 on a damp filter for 128 set of seeds.

No. of germinations	0	1	2	3	4	5	6	7	Total
No. of samples	7	6	19	35	30	23	7	1	128

Fit a binomial distribution and find the expected frequencies if the chance of seed germination is 0.5

Also, find mean and standard deviation for the same.

Binomial Fitting

p	q	N
0.5	0.5	128

Since there are 8 terms, therefore, $n = 7$

Expected frequencies are nothing but coefficients of binomial expansion of $N(p+q)^n$. Hence, a generalised formula for expected frequency is-

$$f_x = {}^nC_x$$

Calculating mean, variance and standard deviation

- Mean
  Mean = np = 7 x 0.5 = 3.5

- Variance
  Variance = npq = 7 x 0.5 x 0.5 = 1.75

- Standard Deviation

Poisson Distribution

HISTORICAL NOTE

Discovered by Mathematician Simeon Poisson in France in 1781

The modeling distribution that takes his name was originally derived as an approximation to the binomial distribution.

POISSON EXPERIMENT PROPERTIES

- Counting the number of times a success occur in an interval
- Probability of success the same for all intervals of equal size
- Number of successes in interval independent of number of successes in other intervals
- Probability of success is proportional to the size of the interval
- Intervals do not overlap

THE POISSON EXPERIMENT

- Used to find the probability of a rare event
- Randomly occurring in a specified interval
  - Time
  - Distance
  - Volume
- Measure number of rare events occurred in the specified interval

ROLE OF POISSON DISTRIBUTION

- It is used in quality control statistics to count the numbers of defects of any item
- In biology to count the number of bacteria
- In physics to count the number of particles emitted from a radio active substance
- In insurance problems to count the number of casualities
- In waiting time problems to count the number of incoming telephone number of incoming telephone calls or incoming customers
- The number of typographical errors per page in typed material

Calculating the Poisson Probability

Determining x successes in the interval

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

μ = mean number of successes in interval
e = base of natural logarithm = 2.71828

The Poisson Experiment Example

On average the anti-virus program detects 2 viruses per week on a notebook

- Time interval of one week

Are the conditions required for the Poisson experiment met?

- $\mu = 2$ per week
- Occurrence of viruses are independent
- yes we Can calculate the probabilities of a certain number of viruses in the interval

Poisson Experiment Example

Let X be the Poisson random variable indicating the number of viruses found on a notebook per week

$P (X = x) = P (x) = \mu^x e^{-\mu} /x!$

$P (x = 0) = P (0) = 2^0 e^{-2} /0! = 0.1353$

$P (x = 1) = P (1) = 2^1 e^{-2} /1! = 0.2707$

$P (x = 3) = P (3) = 2^3 e^{-2} /3! = 0.1804$

$P (x = 7) = P (7) = 2^7 e^{-2} /7! = 0.0034$

X	P (x)
0	0.1353
1	0.2707
2	0.2707
3	0.1804
$\downarrow$	$\downarrow$
7	0.0034
	$\sum p (x) = 1$

The Poisson Experiment Example

Calculate the probability that two or less than two viruses will be found per week

$P (x \leq 2)$

$= p (x = 0) + p (x =1) = p (x = 2)$

$= 0.1353 + 0.2707 + 0.2707$

$= 0.6767$

The Poisson Experiment Example

Calculate the probability that less than three viruses will be found per week

$P (x < 3)$

$= p (x = 0) + p (x =1) = p (x = 2)$

$= 0.1353 + 0.2707 + 0.2707$

$= 0.6767$

The Poisson Experiment Example

Calculate the probability that more than three viruses will be found per week

$P (x > 3)$

$= p (x=4) + p (x = 5) + \ldots\ldots\ldots$

$= 1- p (x \leq 3)$

$1 – 0.8571$

0.1429

The Poisson Experiment Example

Calculate the probability that four viruses will be found in four weeks

$\mu = 2 \times 4 = 8$ in four weeks

$P(x = 4)$

$= 8^4 e^{-8} / 4!$

$= 0.0573$

The Poisson Experiment Example

- Mean and standard deviation of Poisson random variable

$$\mu = E(x) = \mu$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{Var(X)} = \sqrt{\mu}$$

Fitting a Poisson distribution

The following tables gives a number of days in a 50 – day period during which automobile accidents occurred in a city

Fit a Poisson distribution to the data :

No of accidents : 0 1 2 3 4

No. of days :21 18 7 3 1

solution

Fitting of Poisson distribution

X	F	Fx
0	21	0
1	18	18
2	7	14
3	3	9
4	1	4
	N=50	∑fx=45

$M = \bar{X} = \sum fx/N = 45/50 = 0.9$

Fitting a Poisson distribution

$f(0) = e^{-m} = e^{-0.9} = 0.4066$

$f(1) = mf(0) = (0.9)(0.4066) = 0.3659$

$f(2) = m/2\ f(1) = 0.9/2\ (0.3659) = 0.1647$

$f(3) = m/3\ f(2) = 0.9/3\ (0.1647) = 0.0494$

$f(4) = m/4\ f(3) = 0.9/4\ (0.0494) = 0.0111$

In order to fit Poisson distribution , we shall multiply each probability by N i.e. 50

Hence the expected frequencies are :

X	0	1	2	3	4
F	0.4066 X 50	0.3659 X 50	0.1647 X 50	0.0494 X 50	0.0111 X 50
	=20.33	=18.30	=8.24	=2.47	=0.56

Thank You