Discrete Probability Distributions

Prof. dr. Siswanto Agus Wilopo, M.Sc., Sc.D.
Department of Biostatistics, Epidemiology and Population Health
Faculty of Medicine
Universitas Gadjah Mada

sawilopo@yahoo.com Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Learning Objectives

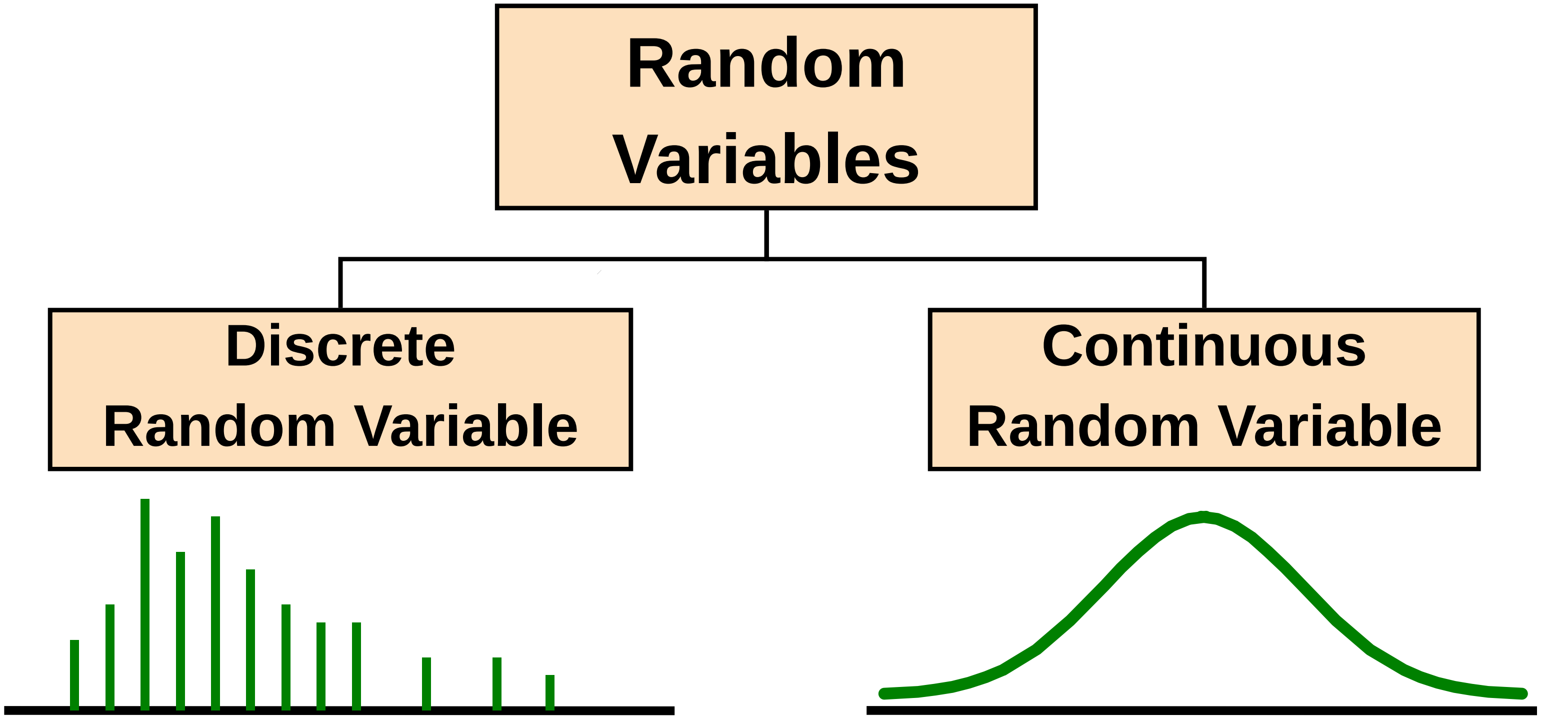
In this lecture, you learn:

- The properties of a probability distribution
- To calculate the expected value, variance, and standard deviation of a probability distribution
- To calculate probabilities from binomial and Poisson distributions
- How to use the binomial and Poisson distributions to solve business problems

sawilopo@yahoo.com Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Introduction to Probability Distributions

- **Random Variable**
  - **Represents a possible numerical value from an uncertain event**

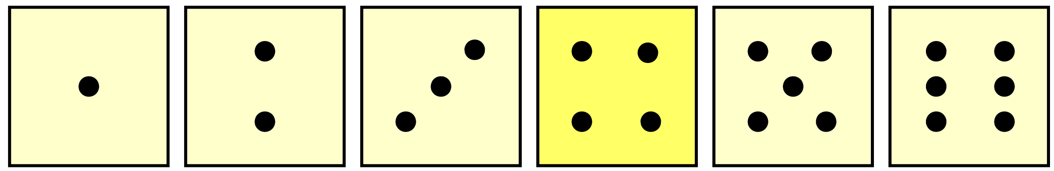


sawilopo@yahoo.com

Discrete Random Variables

- Can only assume a countable number of values

Examples:



- Roll a die twice
  Let X be the number of times 4 comes up
  (then X could be 0, 1, or 2 times)

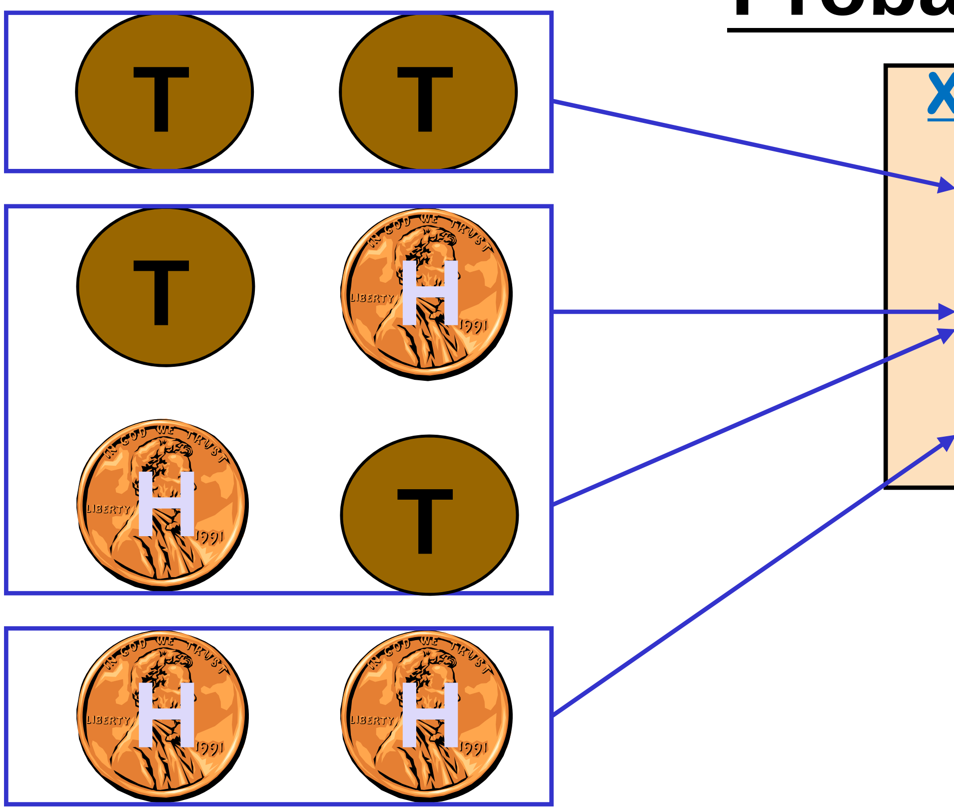
- Toss a coin 5 times.
  Let X be the number of heads
  (then X = 0, 1, 2, 3, 4, or 5)

sawilopo@yahoo.com
Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Discrete Probability Distribution

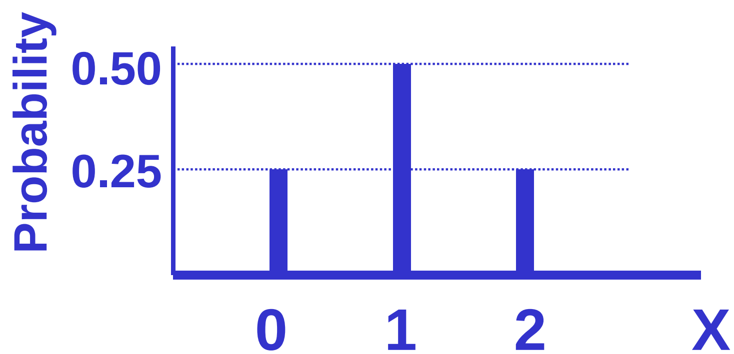
Experiment: Toss 2 Coins. Let X = # heads.

4 possible outcomes



Probability Distribution

X Value	Probability
0	1/4 = 0.25
1	2/4 = 0.50
2	1/4 = 0.25



Discrete Random Variable Summary Measures

- Expected Value (or mean) of a discrete distribution (Weighted Average)

$$\mu = E(X) = \sum_{i=1}^{N} X_i P(X_i)$$

- Example: Toss 2 coins,
  X = # of heads,
  compute expected value of X:

E(X) = (0 x 0.25) + (1 x 0.50) + (2 x 0.25)
= 1.0

X	P(X)
0	0.25
1	0.50
2	0.25

sawilopo@yahoo.com
Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Discrete Random Variable Summary Measures

(continued)

- Variance of a discrete random variable

$$\sigma^2 = \sum_{i=1}^{N} [X_i - E(X)]^2 P(X_i)$$

- Standard Deviation of a discrete random variable

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^{N} [X_i - E(X)]^2 P(X_i)}$$

where:

E(X) = Expected value of the discrete random variable X

X_i = the i^{th} outcome of X

$P(X_i)$ = Probability of the i^{th} occurrence of X

sawilopo@yahoo.com

Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Discrete Random Variable Summary Measures

(continued)

- **Example: Toss 2 coins, X = # heads, compute standard deviation (recall E(X) = 1)**

$$\sigma = \sqrt{\sum [X_i - E(X)]^2 P(X_i)}$$

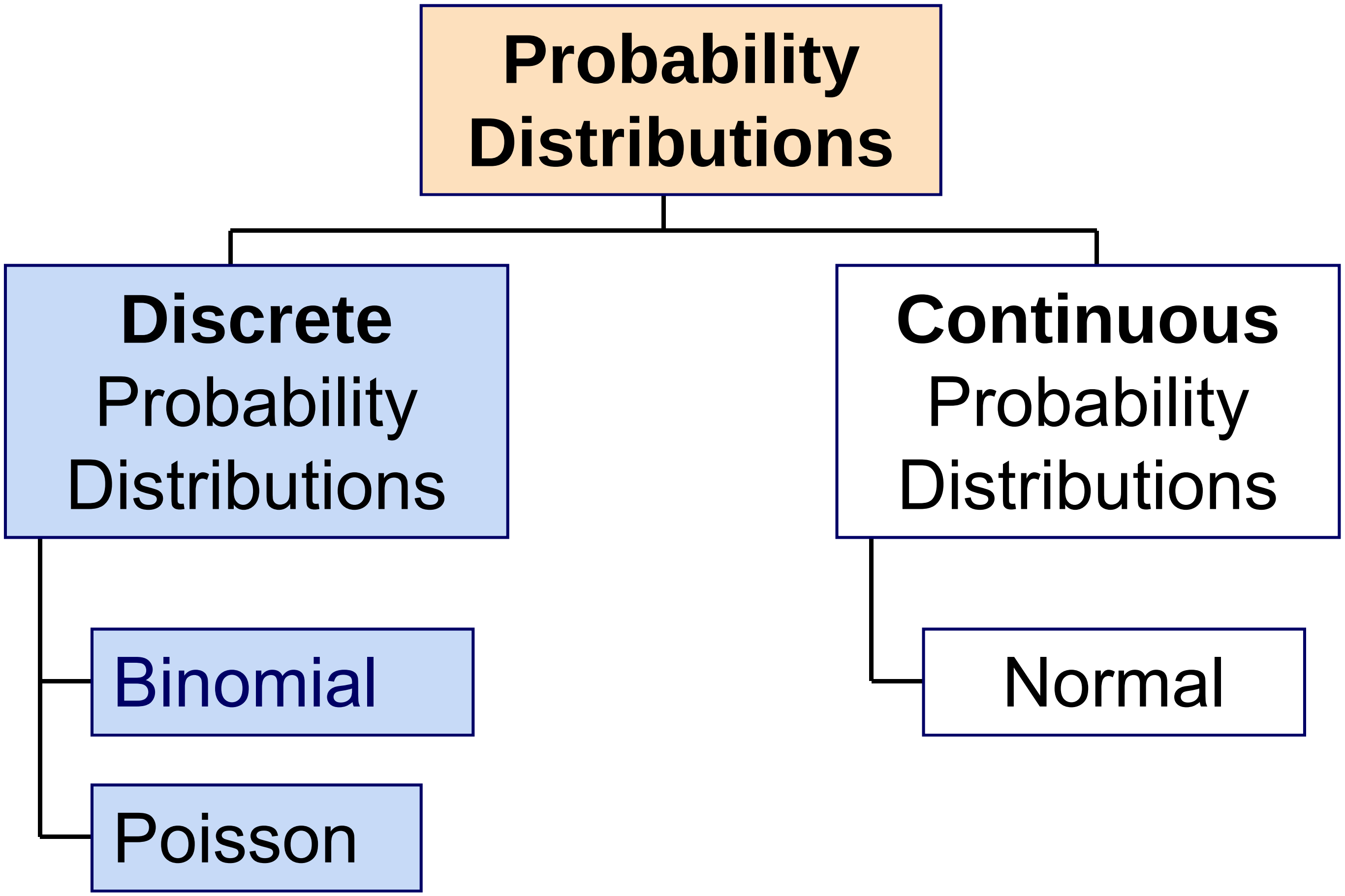
$$\sigma = \sqrt{(0-1)^2(0.25) + (1-1)^2(0.50) + (2-1)^2(0.25)} = \sqrt{0.50} = 0.707$$

Possible number of heads = 0, 1, or 2

sawilopo@yahoo.com

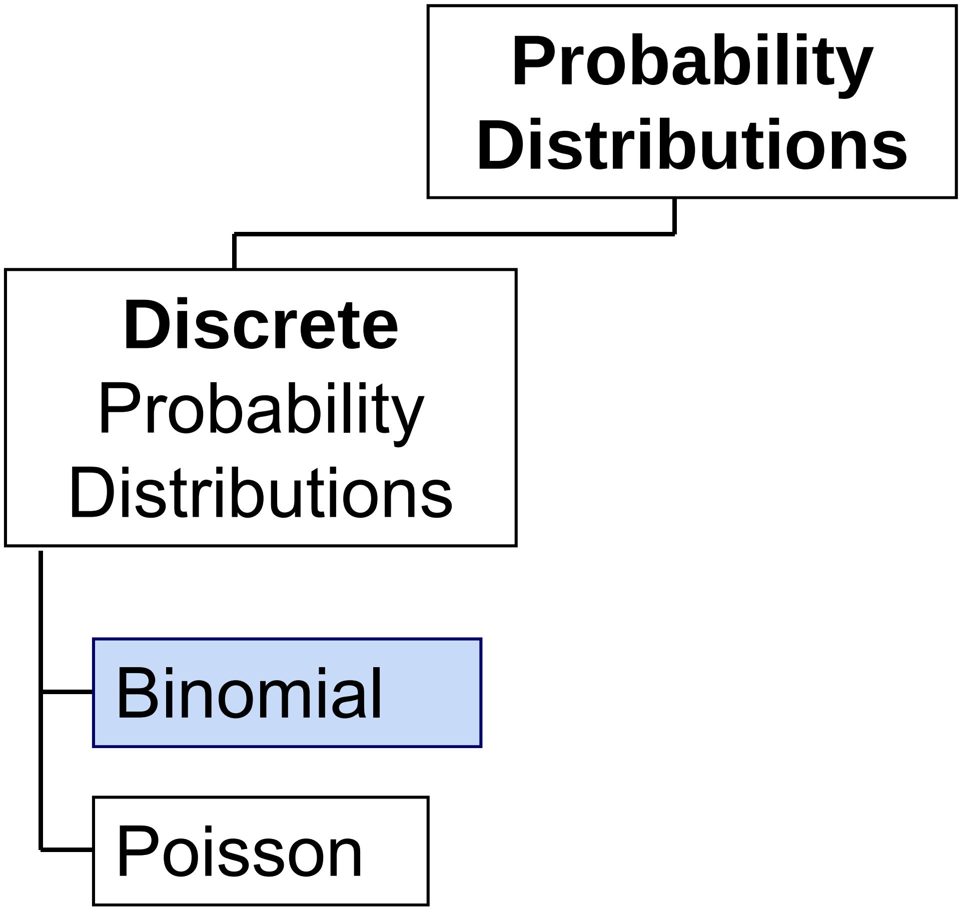
Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Probability Distributions



sawilopo@yahoo.com
Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

The Binomial Distribution



sawilopo@yahoo.com
Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Binomial Probability Distribution

- A fixed number of observations, n
  - e.g., 15 tosses of a coin; ten light bulbs taken from a warehouse
- Two mutually exclusive and collectively exhaustive categories
  - e.g., head or tail in each toss of a coin; defective or not defective light bulb
  - Generally called “success” and “failure”
  - Probability of success is p, probability of failure is 1 – p
- Constant probability for each observation
  - e.g., Probability of getting a tail is the same each time we toss the coin

sawilopo@yahoo.com

Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Binomial Probability Distribution

(continued)

- **Observations are independent**
  - The outcome of one observation does not affect the outcome of the other
- **Two sampling methods**
  - Infinite population without replacement
  - Finite population with replacement

sawilopo@yahoo.com

Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Possible Binomial Distribution Settings

- A patient either cured or died after treatment
- A manufacturing plant labels items as either defective or acceptable
- A firm bidding for contracts will either get a contract or not
- A marketing research firm receives survey responses of "yes I will buy" or "no I will not"
- New job applicants either accept the offer or reject it

sawilopo@yahoo.com

Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Rule of Combinations

- The number of combinations of selecting X objects out of n objects is

$$ {}_{n}C_{x} = \frac{n!}{X!(n-X)!} $$

where:

$$ n! = (n)(n-1)(n-2)\ldots(2)(1) $$

$$ X! = (X)(X-1)(X-2)\ldots(2)(1) $$

$$ 0! = 1 \quad \text{(by definition)} $$

sawilopo@yahoo.com

Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Binomial Distribution Formula

$$P(X) = \frac{n!}{X!(n-X)!} p^X (1-p)^{n-X}$$

P(X) = probability of **X** successes in **n** trials, with probability of success **p** on each trial

X = number of 'successes' in sample, (X = 0, 1, 2, ..., n)

n = sample size (number of trials or observations)

p = probability of "success"

Example: Flip a coin four times, let x = # heads:

n = 4

p = 0.5

1 - p = (1 - 0.5) = 0.5

X = 0, 1, 2, 3, 4

sawilopo@yahoo.com

Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Example:
Calculating a Binomial Probability

What is the probability of one success in five observations if the probability of success is .1?

X = 1, n = 5, and p = 0.1

$$
\begin{aligned}
P(X=1) &= \frac{n!}{X!(n-X)!}p^X(1-p)^{n-X} \\
&= \frac{5!}{1!(5-1)!}(0.1)^1(1-0.1)^{5-1} \\
&= (5)(0.1)(0.9)^4 \\
&= 0.32805
\end{aligned}
$$

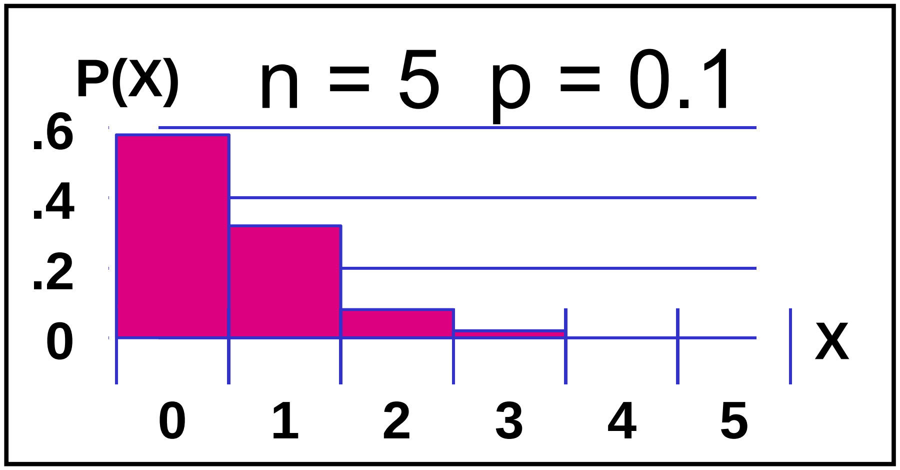
sawilopo@yahoo.com

Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

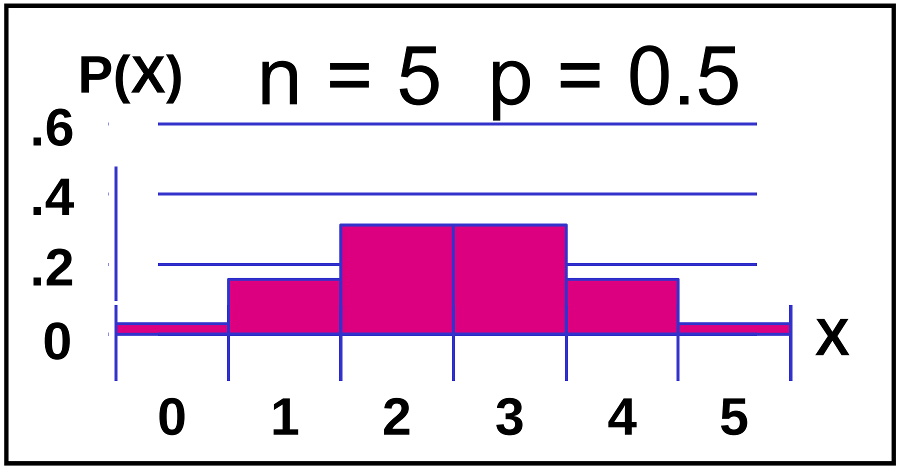
Binomial Distribution

- The shape of the binomial distribution depends on the values of p and n

- Here, n = 5 and p = 0.1



- Here, n = 5 and p = 0.5



Binomial Distribution Characteristics

- **Mean**

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

- Variance and Standard Deviation

$$\sigma^2 = np(1-p)$$

$$\sigma = \sqrt{np(1-p)}$$

Where n = sample size
p = probability of success
(1 – p) = probability of failure

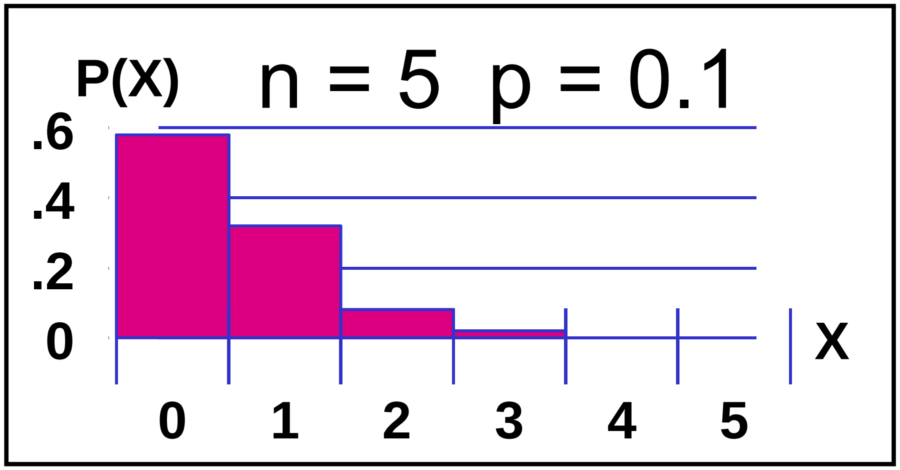
sawilopo@yahoo.com
Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Binomial Characteristics

Examples

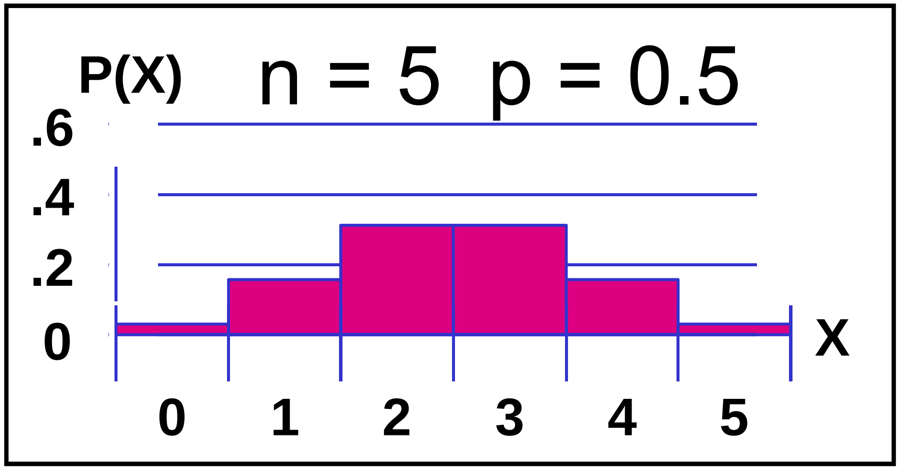
$$\mu = np = (5)(0.1) = 0.5$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{(5)(0.1)(1-0.1)} = 0.6708$$



$$\mu = np = (5)(0.5) = 2.5$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{(5)(0.5)(1-0.5)} = 1.118$$



Using Binomial Tables

n = 10									
x	…	p=.20	p=.25	p=.30	**p=.35**	p=.40	p=.45	p=.50	
0	…	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010	10
1	…	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098	9
2	…	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439	8
3	…	0.2013	0.2503	0.2668	**0.2522**	0.2150	0.1665	0.1172	7
4	…	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051	6
5	…	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461	5
6	…	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051	4
7	…	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172	3
8	…	0.0001	**0.0004**	0.0014	0.0043	0.0106	0.0229	0.0439	**2**
9	…	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098	1
10	…	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0
	…	p=.80	**p=.75**	p=.70	p=.65	p=.60	p=.55	p=.50	**x**

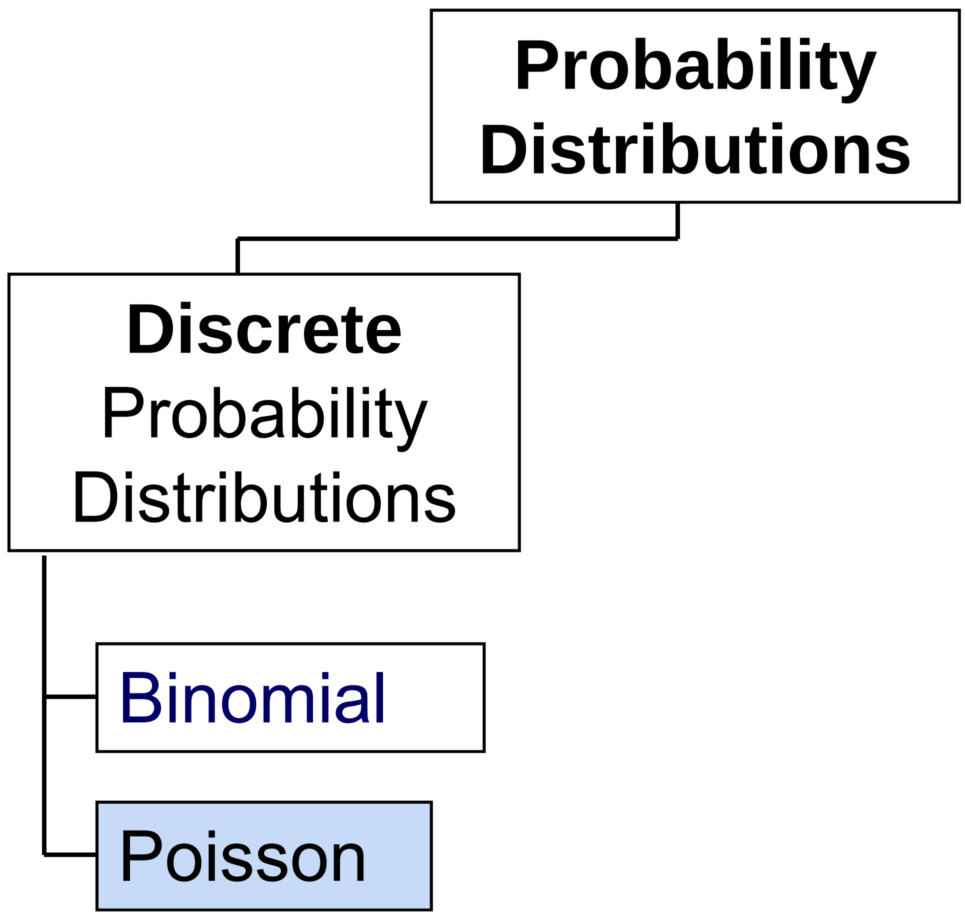
Examples:

n = 10, p = 0.35, x = 3: P(x = 3|n =10, p = 0.35) = 0.2522

n = 10, p = 0.75, x = 2: P(x = 2|n =10, p = 0.75) = 0.0004

sawilopo@yahoo.com Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

The Poisson Distribution



sawilopo@yahoo.com
Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

The Poisson Distribution

- **Apply the Poisson Distribution when:**
  - **You wish to count the number of times an event occurs in a given area of opportunity**
  - **The probability that an event occurs in one area of opportunity is the same for all areas of opportunity**
  - **The number of events that occur in one area of opportunity is independent of the number of events that occur in the other areas of opportunity**
  - **The probability that two or more events occur in an area of opportunity approaches zero as the area of opportunity becomes smaller**
  - **The average number of events per unit is λ (lambda)**

sawilopo@yahoo.com
Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Poisson Distribution Formula

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

where:

X = number of events in an area of opportunity

λ = expected number of events

e = base of the natural logarithm system (2.71828...)

sawilopo@yahoo.com

Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Poisson Distribution Characteristics

- **Mean**

$$\mu = \lambda$$

- Variance and Standard Deviation

$$\sigma^2 = \lambda$$

$$\sigma = \sqrt{\lambda}$$

where λ = expected number of events

sawilopo@yahoo.com Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Using Poisson Tables

X	λ								
	0.10	0.20	0.30	0.40	**0.50**	0.60	0.70	0.80	0.90
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659
2	0.0045	0.0164	0.0333	0.0536	**0.0758**	0.0988	0.1217	0.1438	0.1647
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494
4	0.0000	0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111
5	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0012	0.0020
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Example: Find P(X = 2) if $\lambda = 0.50$

$$P(X=2)=\frac{e^{-\lambda}\lambda^{X}}{X!}=\frac{e^{-0.50}(0.50)^{2}}{2!}=0.0758$$

sawilopo@yahoo.com

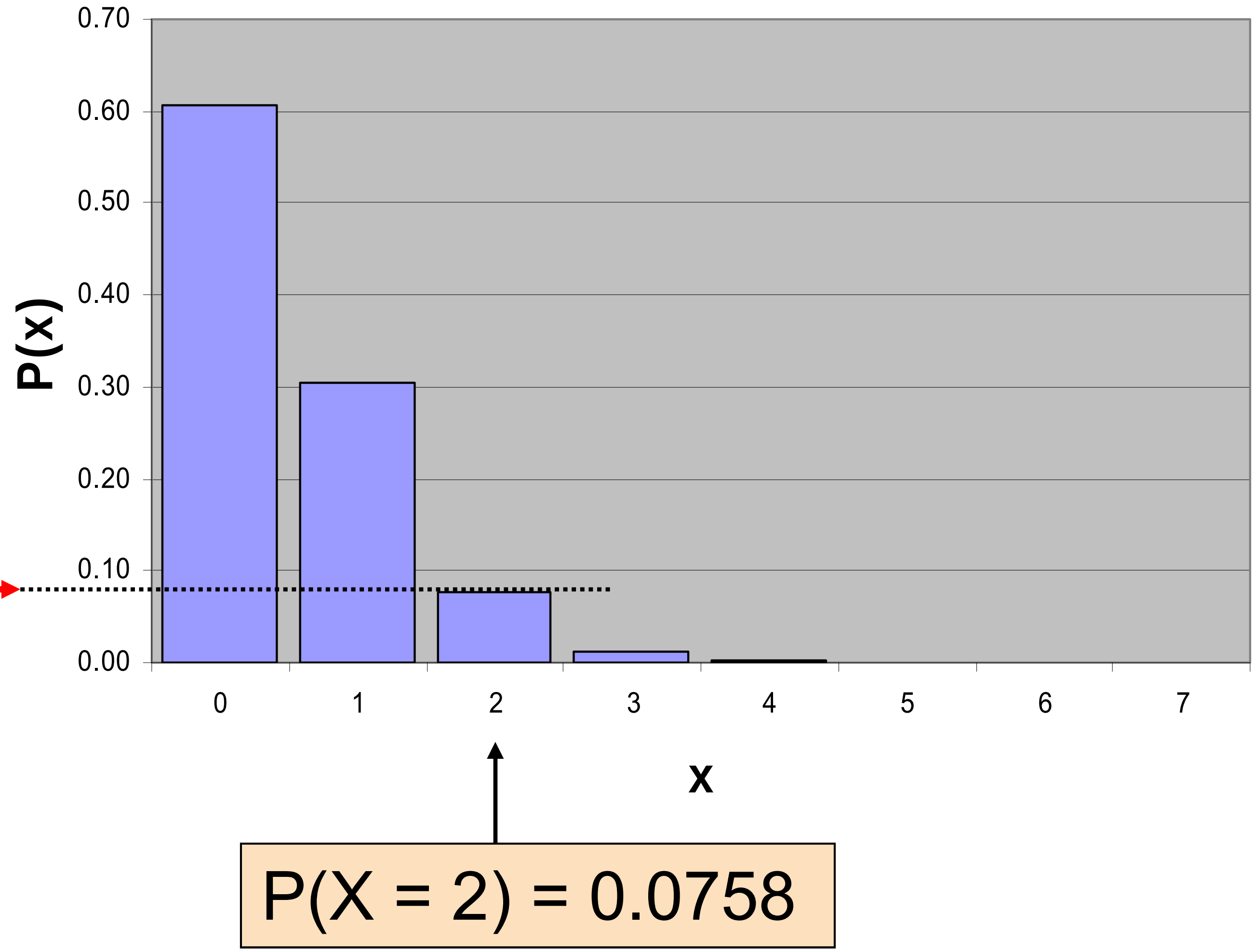
Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Graph of Poisson Probabilities

Graphically:

λ = 0.50

X	λ = 0.50
0	0.6065
1	0.3033
2	0.0758
3	0.0126
4	0.0016
5	0.0002
6	0.0000
7	0.0000

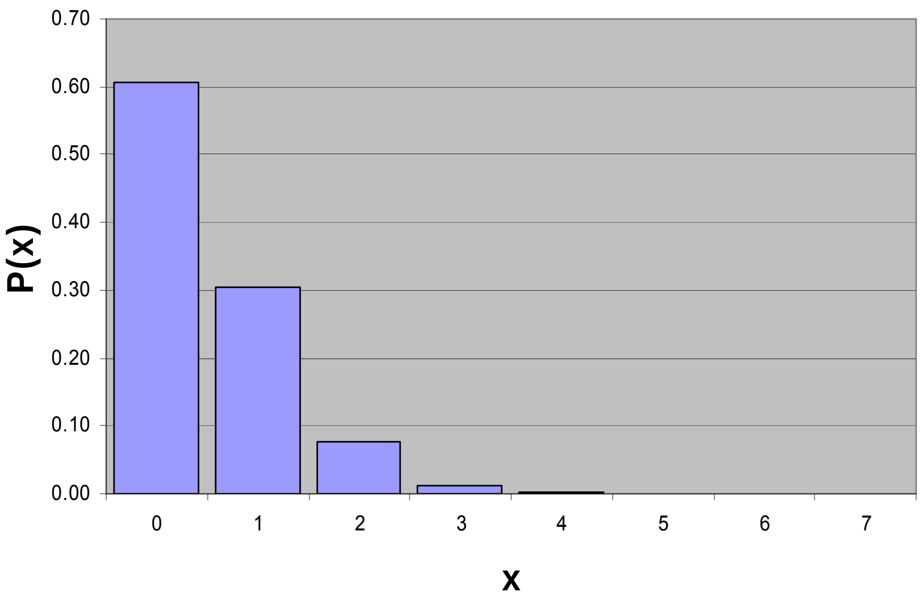


sawilopo@yahoo.com Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

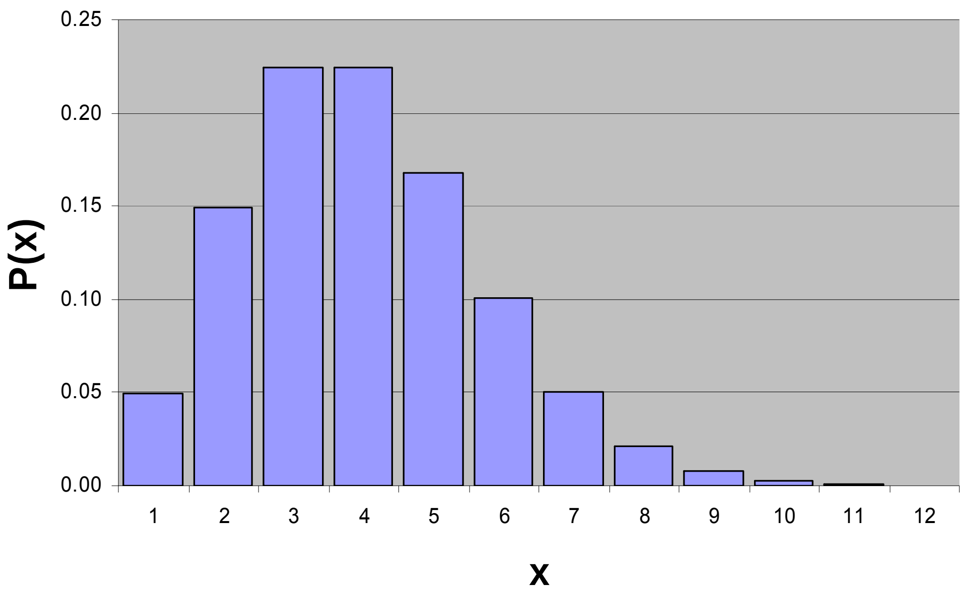
Poisson Distribution Shape

- The shape of the Poisson Distribution depends on the parameter λ :

$\lambda = 0.50$



$\lambda = 3.00$



sawilopo@yahoo.com Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health

Chapter Summary

- Addressed the probability of a discrete random variable
- Discussed the Binomial distribution
- Discussed the Poisson distribution

sawilopo@yahoo.com
Universitas Gadjah Mada, Faculty of Medicine, Department of Biostatistics, Epidemiology and Popualtion Health