Co-funded by the Erasmus+ Programme of the European Union

Practical Testing of Curriculum (WP 3)
Realisation of the Curriculum (WP 4)
MuEuCAP - On-line learning as part of the BioCEP curriculum

STATISTICS AND ENVIRONMENTAL MODELLING
(BIOCEP-621)

Preliminary Concepts of Statistics & Frequency Distribution (Normal/Binomial/Poisson Distribution)

Dr. Khin Swe Oo
Associate Professor
Department of Zoology
Myeik University
khinsweoo@bmwuni.edu.mm

(Part II)

Frequency Distribution (Normal/Binomial/Poisson Distribution)

Outline

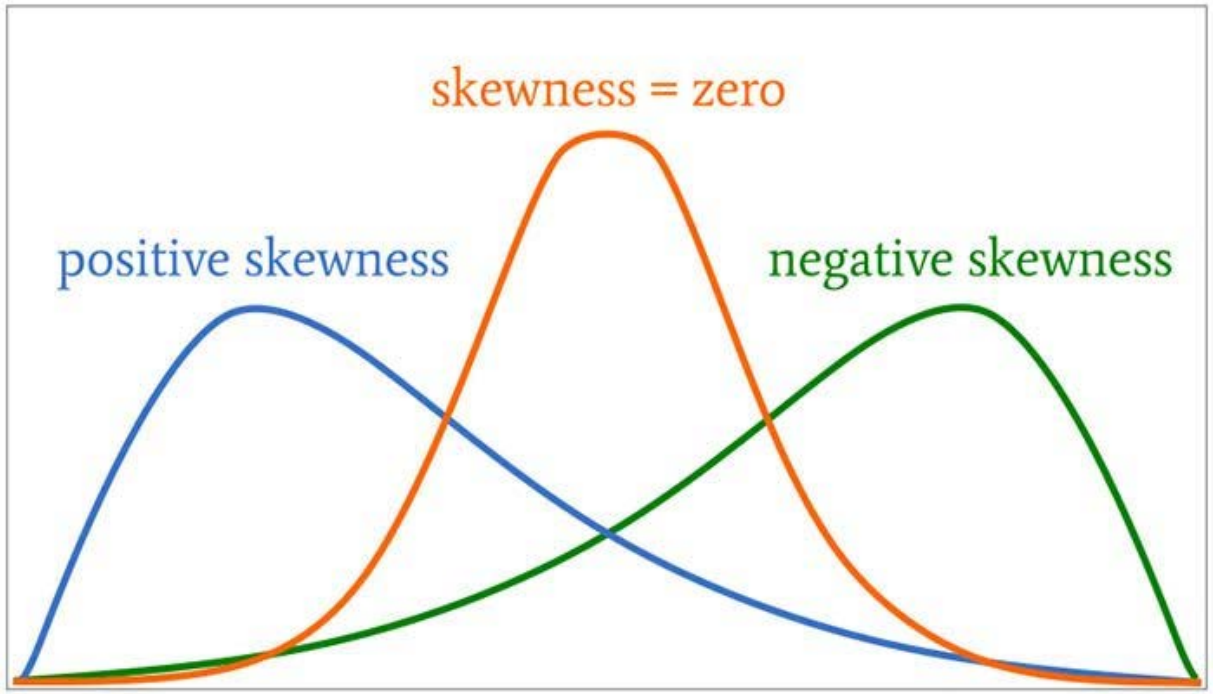
- **Shape of distribution**
  - Skewness and Kurtosis
- **Frequency distribution**
  - Normal distribution
  - Binomial distribution
  - Poisson distribution

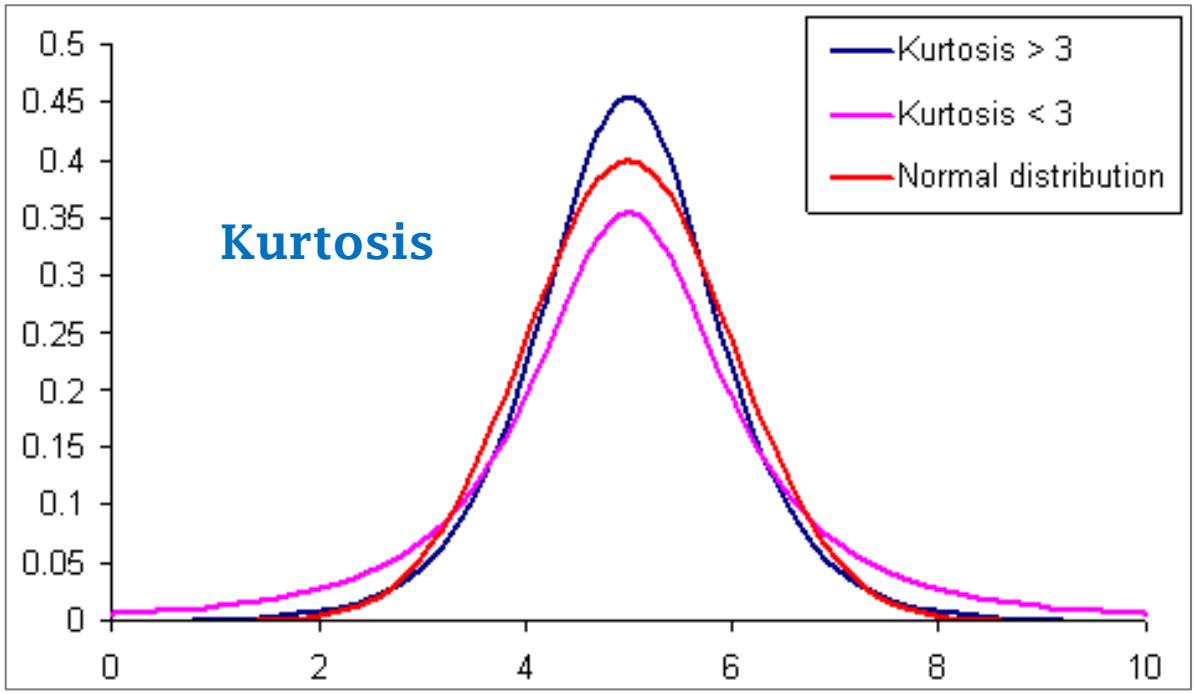
Shape of distribution: Skewness and Kurtosis

Basis for comparison	Skewness	Kurtosis
Meaning	The tendency of a distribution that determines its symmetry about the mean	The measure of the respective sharpness of the curve, in the frequency distribution
Measure for	Degree of distortion in the distribution	Degree of tailedness in the distribution
What is it?	An indicator of lack of equivalence in the frequency distribution	The measure of data, which is either peaked or flat in relation to the normal distribution
represents	how much and in which direction, the values deviate from the mean?	How tall and sharp the central peak is?

Source: https://keydifferences.com/differences-between-skewness-and-kurtosis.html

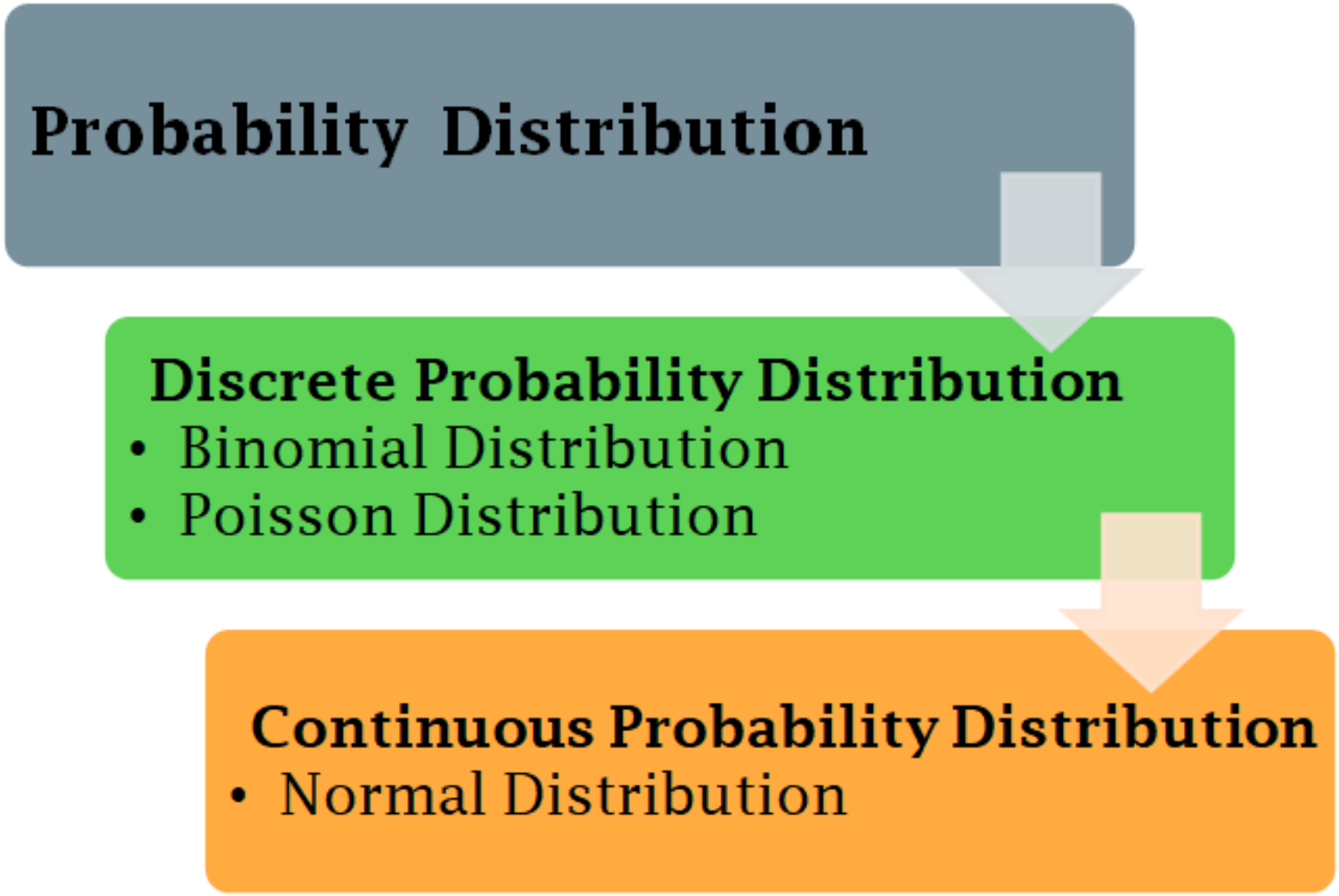
Shape of distribution: Skewness and Kurtosis





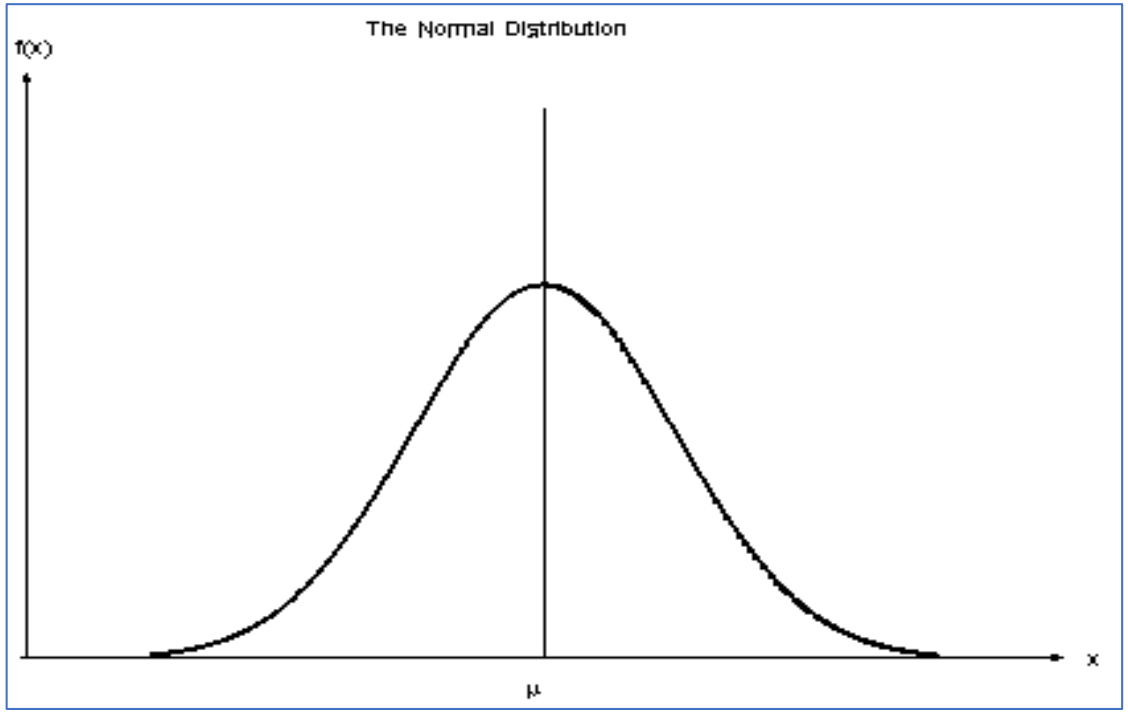
Source: https://www.allaboutcircuits.com/technical-articles/understanding-the-normal-distribution-parametric-tests-skewness-and-kurtosis/

Frequency Distribution (Probability Distribution)



Normal Distribution

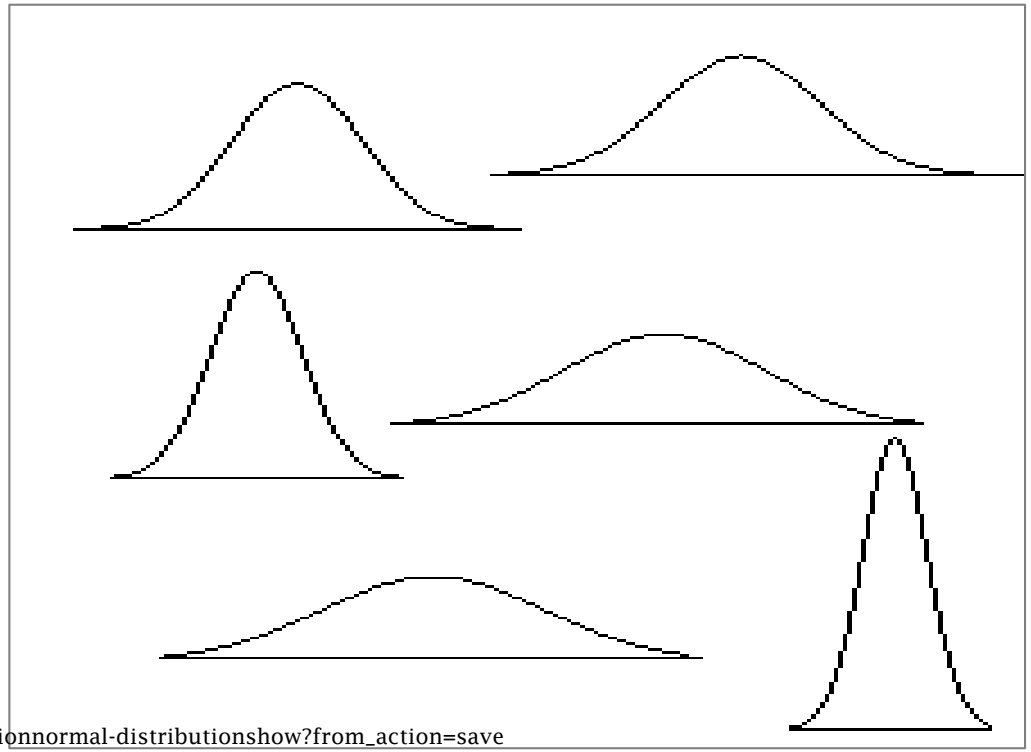
- a symmetrical probability distribution
- most results in the middle and few spread on both sides
- shape of a bell
- described by its mean (μ) and standard deviation (σ)



Source: https://magoosh.com/statistics/understanding-normal-distribution/

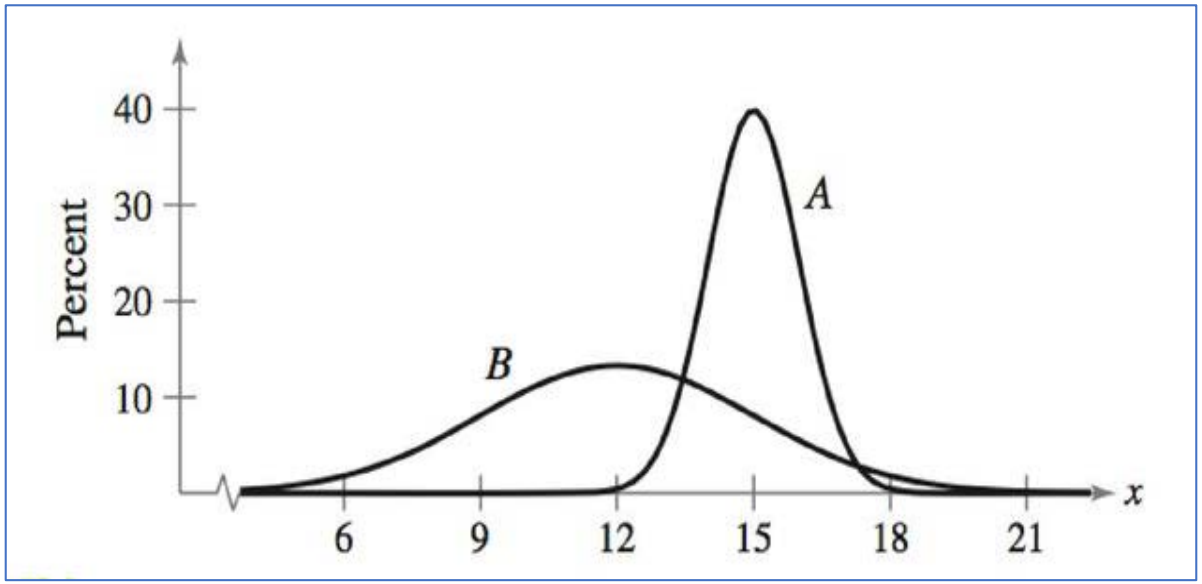
Properties of Normal Distribution

- the mean, median and mode are equal
- the area under the curve is always the same
- concentrated more in the middle than the tails
- the total area under the curve is 1



Source: https://www2.slideshare.net/BiologyIB/standard-deviationnormal-distributionshow?from_action=save

Example. 1: Which normal curve has the greater mean?



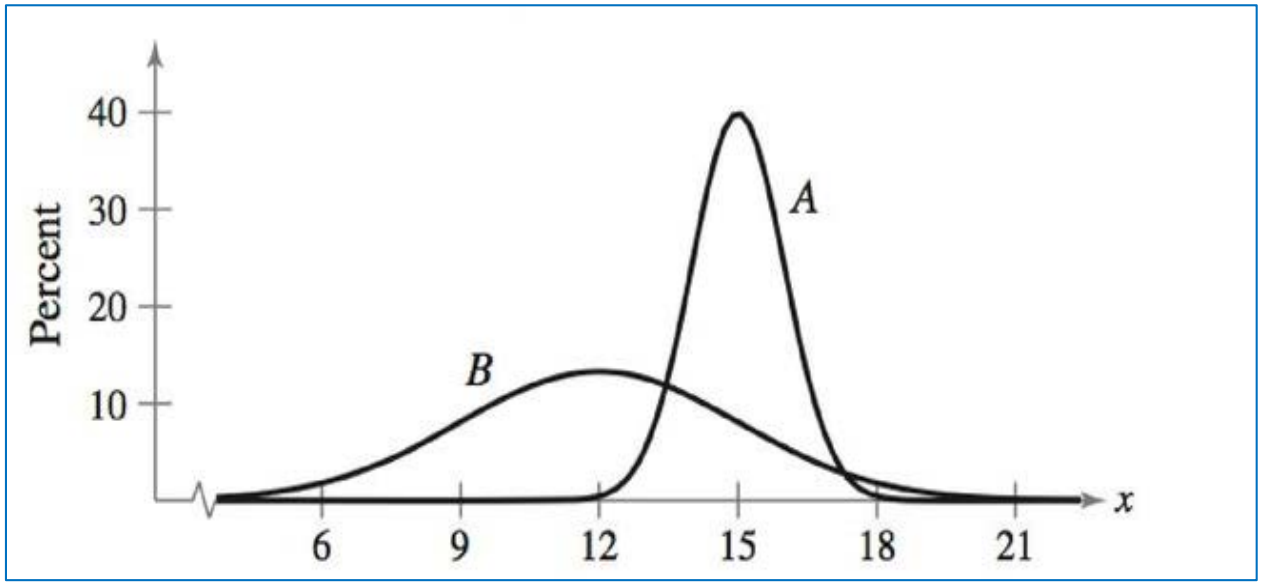
Solution:

The line of symmetry of curve A occurs at x = 15

The line of symmetry of curve B occurs at x = 12

So, Curve A has the greater mean.

Example. 2: Which curve has the greater standard deviation?



Solution:

Curve B is more spread out than Curve A

So, Curve B has the greater standard deviation

Source: Pearson Education,2012

Binomial Distribution

- a discrete probability distribution
- used for data which can only take one of two values, i.e.,
    - Pass or fail
    - Yes or no
    - Good or defective

Properties of Binomial Distribution

- experiment consists of n repeated trials
- an outcome may be classified as a **success** or a **failure**
- the probability of a success, denoted by p, remains constant from trial to trial and repeated trials are independent
- the number of successes X in n trials of a binomial experiments is called a binomial random variable

Properties of Binomial Distribution

- The probability distribution of the random variable ***X*** is called a binomial distribution, and is given by the formula:

$$\mathbf{P\,(X) = \; C^{n}_{x}\, p^{x}\, q^{n\,-\,x}}$$

where

n = the number of trials

x = 0, 1, 2, ... N

p = the probability of success in a single trial (i.e. q = 1 – p)

C^{n}_{x} = a combination

P (X) = the probability of successes in ***n*** binomial trials

Example: A die is tossed 3 times. What is the probability of (a) No fives turning up? (b) 1 five? (c) 3 fives?

Source: pixabay.com

Solution:

n = 3, p = 1/6, q = 5/6

Let X = number of fives appearing

(a) x = 0

$P(X = 0) = \mathbf{C^n_x\ p^x\ q^{n-x}} = C^3_0\,(1/6)^0\,(5/6)^{3-0} = C^3_0\,(1/6)^0\,(5/6)^3$

$= 125/216 = 0.5787$

(b) x = 1

$P(X = 1) = \mathbf{C^n_x\ p^x\ q^{n-x}} = C^3_1\,(1/6)^1\,(5/6)^2 = 75/216 = 0.34722$

(c) X = 3

$P(X = 3) = \mathbf{C^n_x\ p^x\ q^{n-x}} = C^3_3\,(1/6)^3\,(5/6)^0 = 1/216 = 4.6296 \times 10^{-3}$

Poisson Distribution

- a discrete probability distribution for the counts of events that occur randomly in a given interval of time (or space)

- many experimental situations occur in which the counts of events within a set unit of time, area, volume, length, etc.

Properties of Poisson Distribution

- independent of the other occurrences
- the occurrences in each interval range from zero to infinity
- the mean number of occurrences/events must be constant throughout the experiment

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

where, x = 0, 1, 2, 3, ...

λ = mean number of occurrences in the interval

e = Base of the natural logarithm = 2.71828

k = the number of occurrences of an event;

Example. 1: A typist makes on average 2 mistakes per page. What is the probability of a particular page having no errors on it?

Source: secretarytypistwork.com

Solution:

Let, x be a Poisson random variable

$x = 0, \lambda = 2, p = ½$

λt = (2 errors per page * 1 page) = 2

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

P0 = 2^0/0! * exp(-2) = 0.135 = 13.5%

Example. 2: The mean number of faults in a new house is 8.

What is the probability of buying a new house with exactly 1 fault?

Solution:

Here we have an average rate of faults occurring 8 per house

Hence, $\lambda t = (8 \text{ faults/house} * 1 \text{ house}) = 8$

$n = 1$

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

$P1 = 8\text{^}1/1! * \exp(-8) = 0.0027$

Source: pinterest.com

Difference between Binomial and Poisson Distribution

Binomial distribution	Poisson distribution
- biparametric, i.e., featured by two parameters, n and p	- uniparametric, i.e., characterised by a single parameter, λ
- there are only two possible outcomes, i.e., Success and Failure	- there are an unlimited number of possible outcomes
- Mean (np) > Variance (npq)	- Mean = Variance
Eg - Coin tossing experiment	Eg – Printing mistakes/Page of a large book

Assignments

Choose the Exactly ONE correct answer.

1. Skewness of Normal distribution is -----.

   (a) Negative (b) Positive (c) 0 (d) Undefined

2. In a Binomial distribution, the mean and variance are equal.

   (a) True (b) False

3. In a Poisson Distribution, the mean and variance are equal.

   (a) True (b) False

4. The shape of the Normal Curve is --------.

   (a) Flat (b) Circular (c) Bell shaped (d) Spiked

5. In Normal distribution, the highest value of ordinate occurs at -------.

   (a) Mean (b) Variance (c) Extremes

   (d) Same value occurs at all points

Further information

https://www.docsity.com/en/probability-distributions-statistics-lecture-slides/245765/

https://explorable.com/frequency-distribution

https://www.intmath.com/counting-probability/12-binomial-probability-distributions.php

https://statisticsbyjim.com/basics/normal-distribution/

https://mathbitsnotebook.com/Algebra1/StatisticsData/STShapes.html

Forbes et al., 2011. *Statistical Distributions (Fourth Edition)*. A John Wiley & Sons, In., Publication.

Walck, C., 2007. *Hand-book on Statistical Distributions for Experimentalists*. Particle Physics Group. Fysikum, University of Stockholm.

Rosner, B., 2015. *Fundamentals of Biostatistics (Eighth Edition)*. CENGAGE Learning, Harvard University.

Chap, T.LE., 2003. *Introductory Biostatistics*. Wiley-Interscience. A John Wiley & Sons, Publication.

Take home message

- Differences between Skewness and Kurtosis
- Frequency distribution
    - what is it and what is it for
    - continuous: Normal distribution
    - discrete: Binomial and Poisson distribution