Overview of Biostatistics

Course Outlines

- **Basic Algebra**
- **Functions**
- **What do we call 'Statistics'?**
- **Measures of Central Tendency**
  - **Arithmetic Mean**
  - **Median**
  - **Mode**

Course Outlines

- **Measures of Dispersion**
  - **Range**
  - **Mean Deviation**
  - **Variance and Standard Deviation**
- **Hypothesis Testing**
  - **One Population**
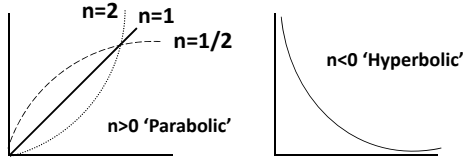 - **Two Populations**
  - **More than two Populations**

Biological Function in Fisheries Biology

- Linear function
- Power function
- Exponential function
- Asymptotic function
- Logistic function
- Parabolic function
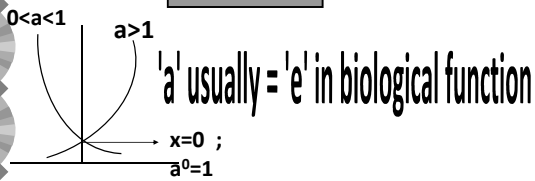- Gompertz function

---

Functions

- Power Function

$$Y = kX^n$$

n=2 n=1
n=1/2

n<0 'Hyperbolic'

n>0 'Parabolic'

Functions

- Logarithmic Function

$$Y = a^X$$

0<a<1   a>1

'a' usually = 'e' in biological function

x=0 ;
$a^0=1$

What do we call 'Statistics'?

- **Definition**
  - **Statistics may be defined as the collection, organization, presentation, analysis and interpretation of numerical data**
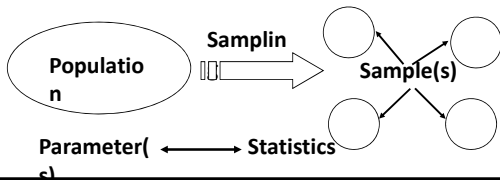- **Types of Statistics**
  - *Descriptive Statistics :* **the phase of statistics which seeks only to describe and analyze a given group without drawing conclusions or inferences about a larger group.**
  - *Inductive or Inference Statistics :* **the phase of statistics which seeks to describe and analyze a given group or sample and draw conclusions or inferences about a larger group.**

What do we call 'Statistics'?

- **Population and Sample**
  - *Population :* **the total set of items defined by a characteristics of these items which we call 'Parameter(s)'**
  - **Sample : a finite set of items drawn from a population, which serves as a representative of such population which we call 'Statistics'**

Population — Samplin(g) → Sample(s)

Parameter(s) ←→ Statistics

What do we call 'Statistics'?

- **Types of Variable**
  - *Continuous Variable :* **a variable which can theoretically assume any value between two given values, e.g. numerical values of length and weight**
  - *Discrete Variable :* **a variable which cannot assume any value between two given values, e.g. sex, no. of fishermen, no. of gears**
- **Grouped and Ungrouped Data**
  - *Ungrouped Data :* **Ordinary data**
  - *Grouped Data :* **Data which we categorized in 'Class Intervals' ; mark ; and examine their 'frequency' in each interval.**

Measures of Central Tendency

- Arithmetic Mean
  - the sum of the set of values divided by their number
  - for ungrouped data :

$$\bar{x} = \frac{\sum x_i}{n}$$

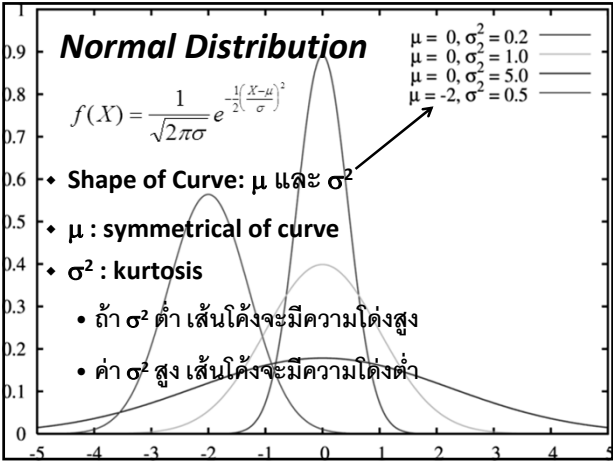 - for grouped data :

$$\bar{x} = \frac{\sum f_i x_i}{\sum f}$$

Measures of Central Tendency
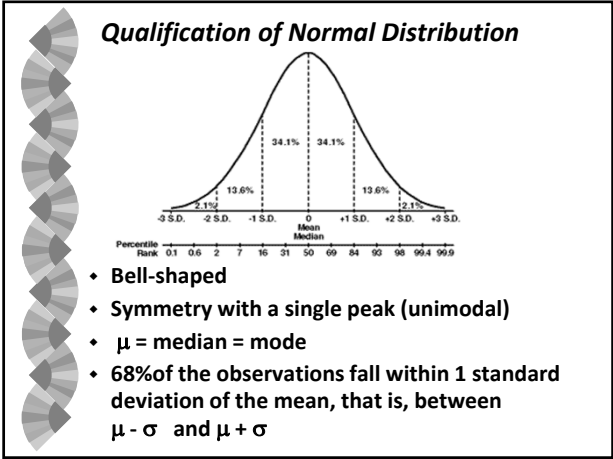
- Median
  - the value of the middle item or the mean of the values of two middle items when the data are arranged in the order of magnitude.
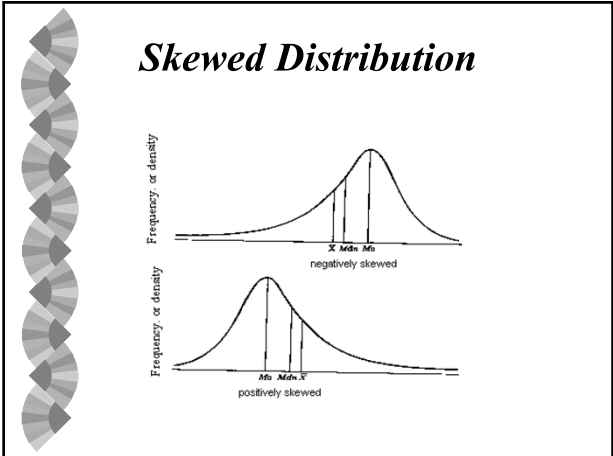
Measures of Central Tendency

- Mode
  - the value which occurs with the highest frequency

Normal Distribution

$$f(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}$$

μ = 0, σ² = 0.2 ——
μ = 0, σ² = 1.0 ——
μ = 0, σ² = 5.0 ——
μ = -2, σ² = 0.5 ——

- **Shape of Curve: μ และ σ²**
- **μ : symmetrical of curve**
- **σ² : kurtosis**
  - **ถ้า σ² ต่ำ เส้นโค้งจะมีความโด่งสูง**
  - **ค่า σ² สูง เส้นโค้งจะมีความโด่งต่ำ**

Qualification of Normal Distribution

34.1% 34.1%

13.6% 13.6%

2.1% 2.1%

-3 S.D. -2 S.D. -1 S.D. 0 Mean Median +1 S.D. +2 S.D. +3 S.D.

Percentile Rank 0.1 0.6 2 16 31 50 69 84 93 98 99.4 99.9

- **Bell-shaped**
- **Symmetry with a single peak (unimodal)**
- **μ = median = mode**
- **68%of the observations fall within 1 standard deviation of the mean, that is, between μ - σ and μ + σ**

Skewed Distribution

Frequency or density

X Mdn Mo
negatively skewed

Frequency or density

Mo Mdn X
positively skewed

Measures of Dispersion

- Range
  - highest value minus the lowest value

Measures of Dispersion

- Variance
  - for grouped data :

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{n\sum x_i^2 - (\sum x_i)^2}{n(n-1)}$$

  - for ungrouped data :

$$s^2 = \frac{\sum f_i(x_i - \bar{x})^2}{\sum f_i - 1} = \frac{\sum f_i \sum f_i x_i^2 - (\sum f_i x_i)^2}{\sum f_i(\sum f_i - 1)}$$

- Standard Deviation

$$s = \sqrt{s^2}$$

5 Steps of Hypothesis Testing

- Create Hypothesis
- Select 'alpha' or confidence level
- Choose the right test statistics and calculation
- Compare with 'Critical Value' or checking the P-value
- Conclusion in 'Human Words'

THE END