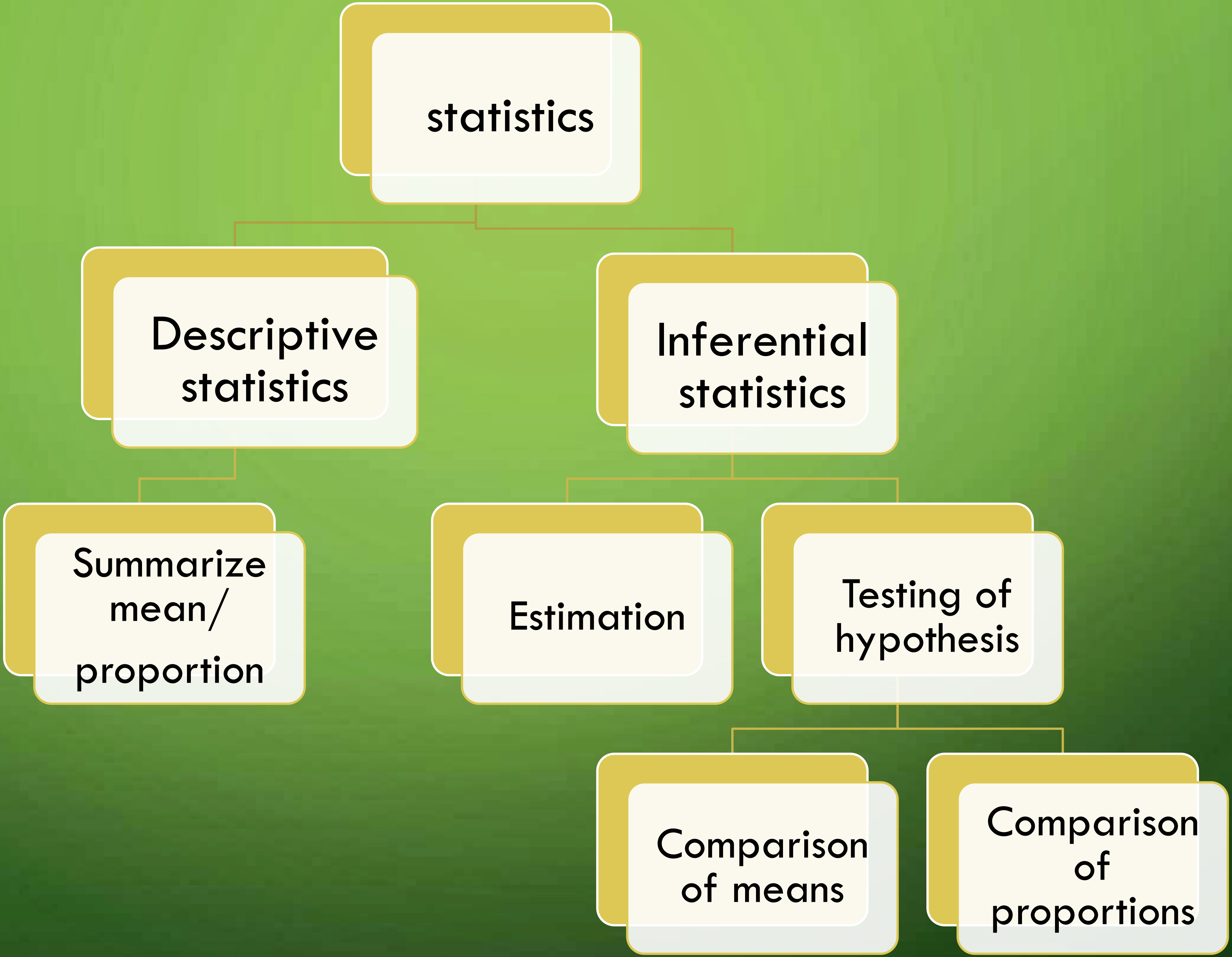
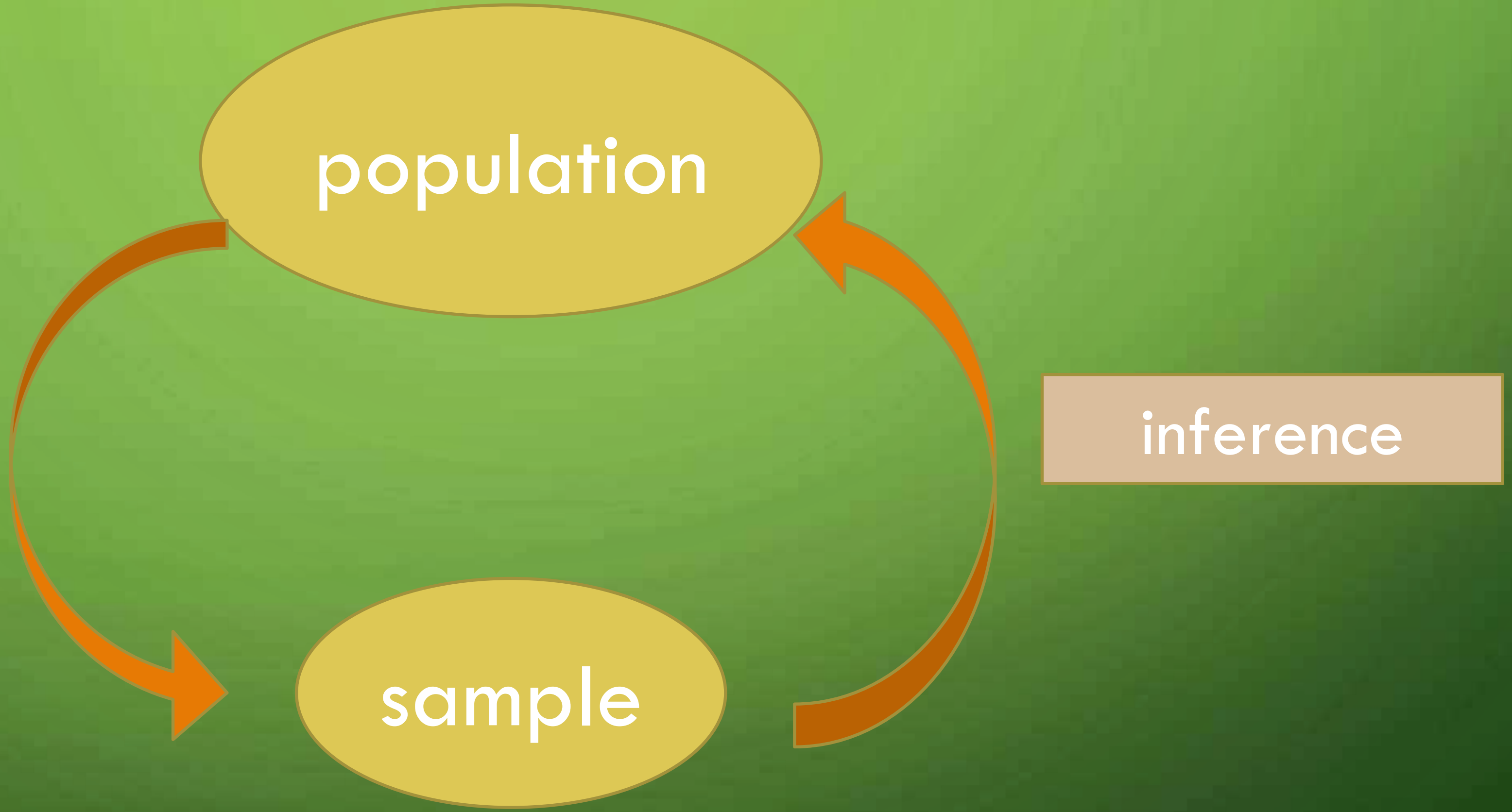
PARAMETRIC TESTS

ONLINE TRAINING WORKSHOP ON BIOSTATISTICS - DAY 4

DR. S. PRAVALLIKA., MD


statistics
Descriptive statistics
Inferential statistics
Summarize mean/ proportion
Estimation
Testing of hypothesis
Comparison of means
Comparison of proportions


population
inference
sample

WHAT IS HYPOTHESIS?

- A statement/ supposition based on observation
- Which may either can be accepted or rejected.

Eg. 1. Average Birth weight of babies is lower in anaemic mothers than non-anaemic mothers.

2. Smokers are at higher risk of developing lung cancer than non-smokers.

TESTING OF HYPOTHESIS

- To test whether the observed difference between two groups or one sample and population value is statistically significant or not.
- Two assumptions/hypotheses are made to draw the inference from sample statistics/value.

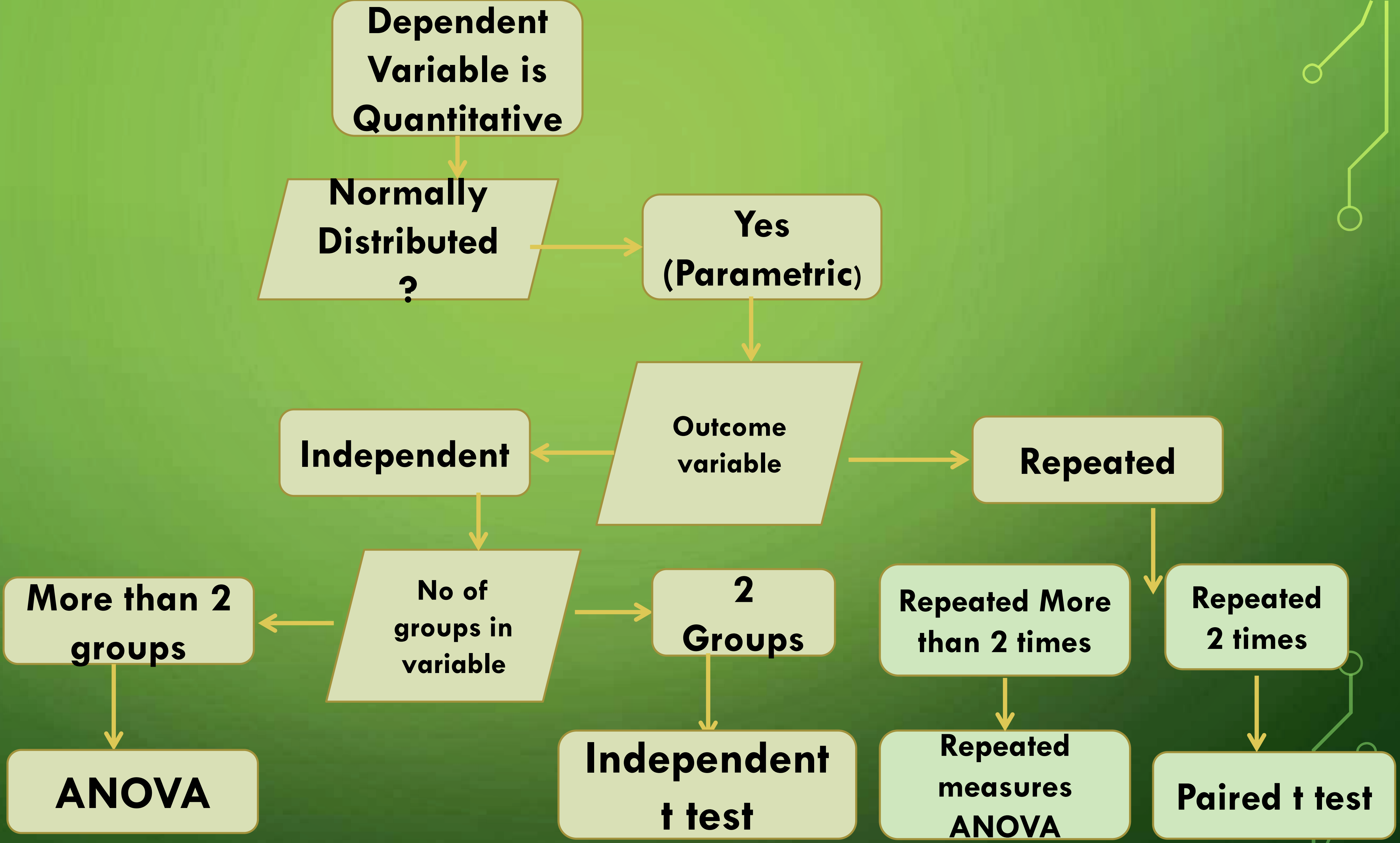
STEPS IN TESTING OF HYPOTHESIS

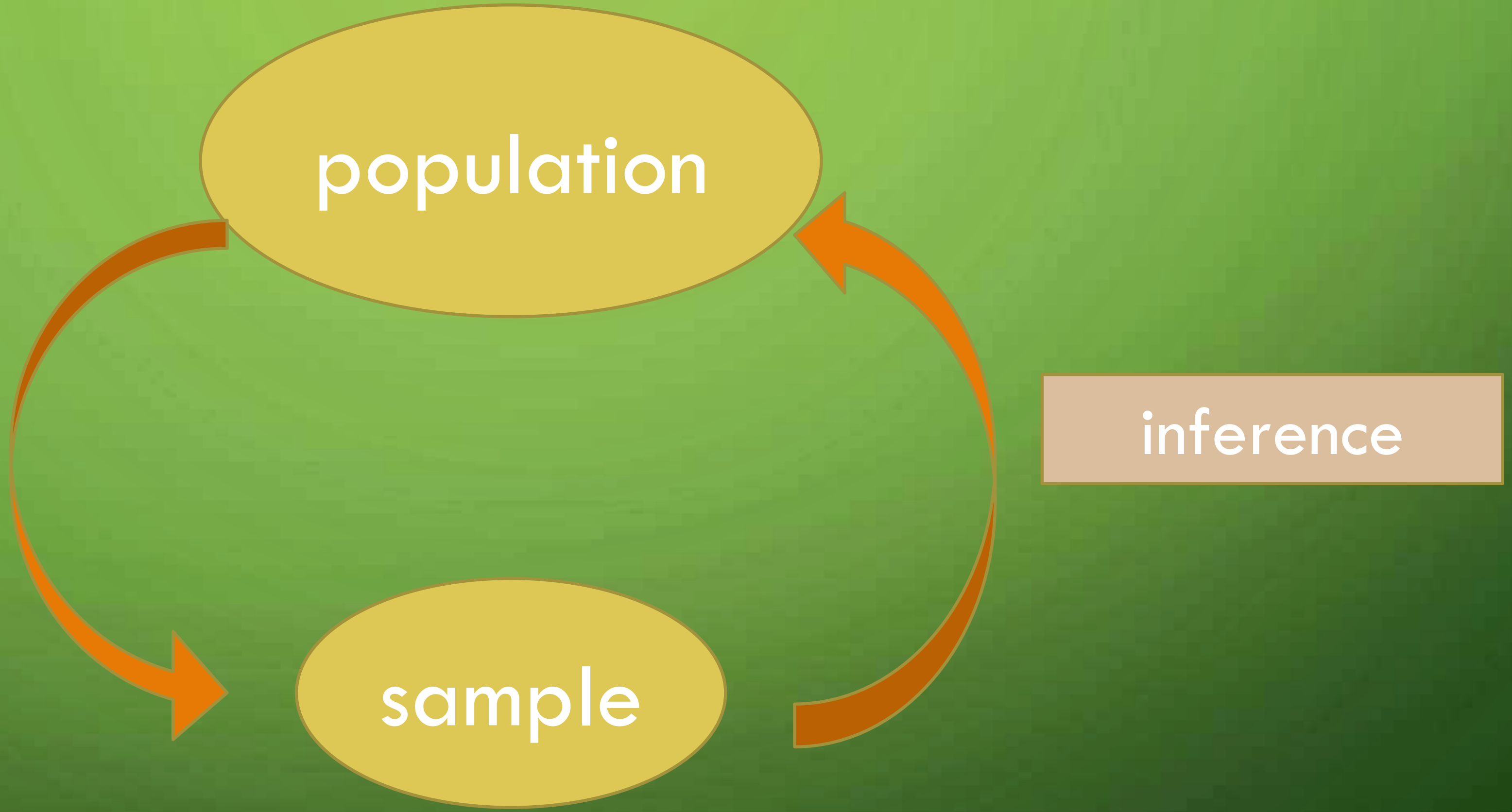
1. Define the variables
2. State the null hypothesis and alternate hypothesis
3. Determine p value (by using appropriate tests of significance)
4. Draw conclusion on the basis of P value.

- Null hypothesis (H_0) or Hypothesis of no difference
- Alternate hypothesis (H1) of significant difference

WHAT IS PARAMETRIC TEST?

- When Data tend to follow one assumed or established distribution (normal, binomial, poisson).
- Test in which population constants (mean, SD, variance) are used.
- In non-parametric tests, no constant of a population is used. Data do not follow any specific distribution. No assumptions are made in non-parametric tests


Dependent Variable is Quantitative
Normally Distributed ?
Yes (Parametric)
Outcome variable
Independent
Repeated
No of groups in variable
More than 2 groups
2 Groups
Repeated More than 2 times
Repeated 2 times
ANOVA
Independent t test
Repeated measures ANOVA
Paired t test

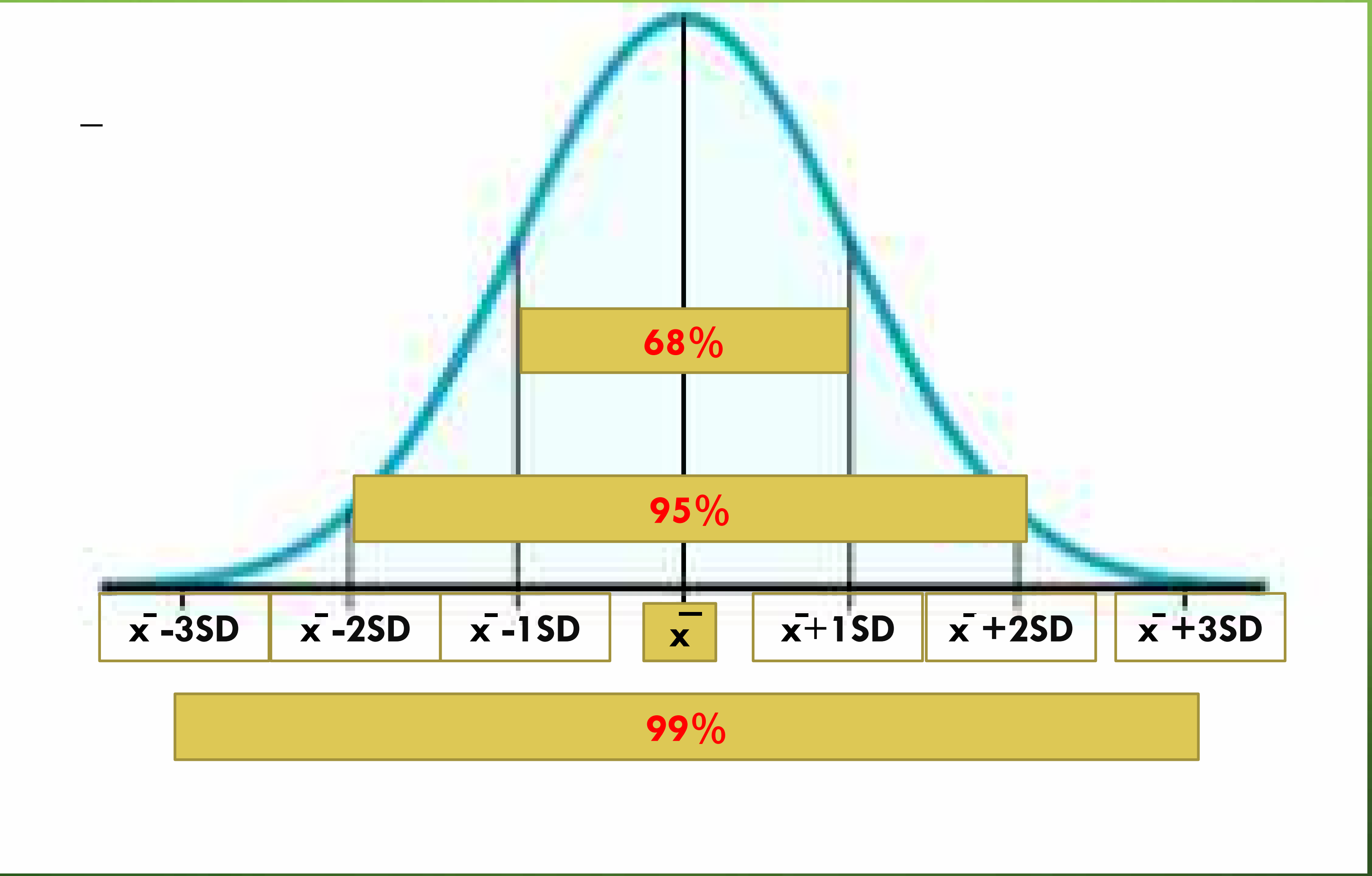

population
sample
inference

STANDARD ERROR (SE)

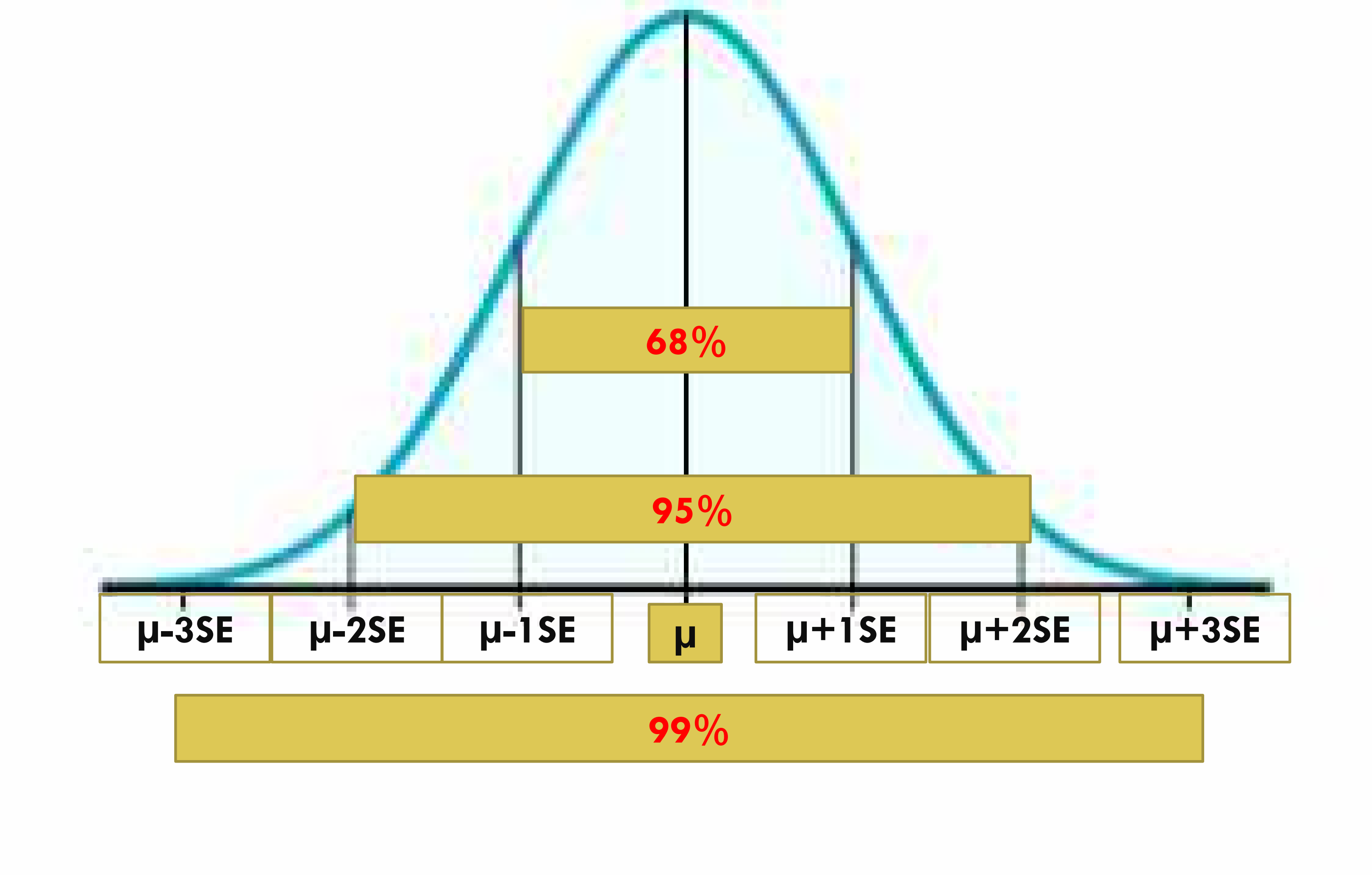
- If we take random samples of same size (n_i) from the population (over and over again), each time we get a different mean. ($\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4 \ldots\ldots \bar{x}_i$)
- If we plot frequency distribution table and histogram with all the sample mean values we get a smooth bell shaped curve with normal distribution.
- Mean of these sample means corresponds to population mean. Dispersion of the sample means around the population mean is measured by standard error (SE).

- Standard deviation is dispersion of individual observations around the sample mean.
- Standard error is dispersion of sample means around population mean.

NORMAL DISTRIBUTION CURVE OF A SAMPLE



NORMAL DISTRIBUTION CURVE OF A POPULATION



- $\mu \pm 1SE$ = 68% sample values
- $\mu \pm 2SE$ = 95% sample values
- $\mu \pm 3SE$ = 99% sample values

- In reality we don't take such no of samples to estimate population mean.
- Standard error (SE)= SD/$\sqrt{n}$
- In estimation of population value,
- 95% Confidence limits = $\bar{x} \pm 2SE$
- 99% Confidence limits = $\bar{x} \pm 3SE$

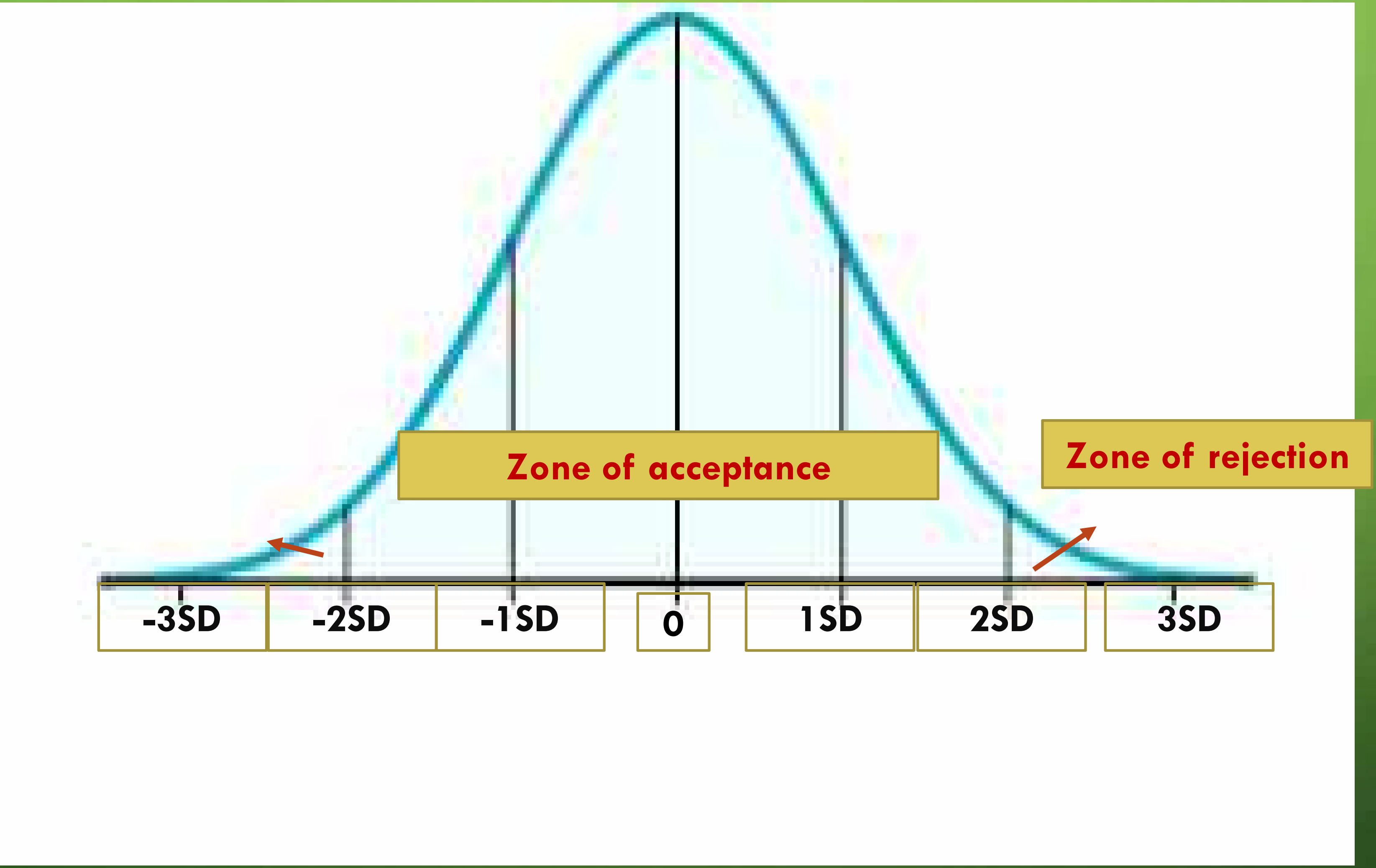
- Standard error of difference between means (SE $_{(\bar{x}_1 - \bar{x}_2)}$) also follows the normal distribution.
- SE $_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{S1^2}{n1} + \frac{S2^2}{n2}}$ (sample size>30)

RELATIVE DEVIATE

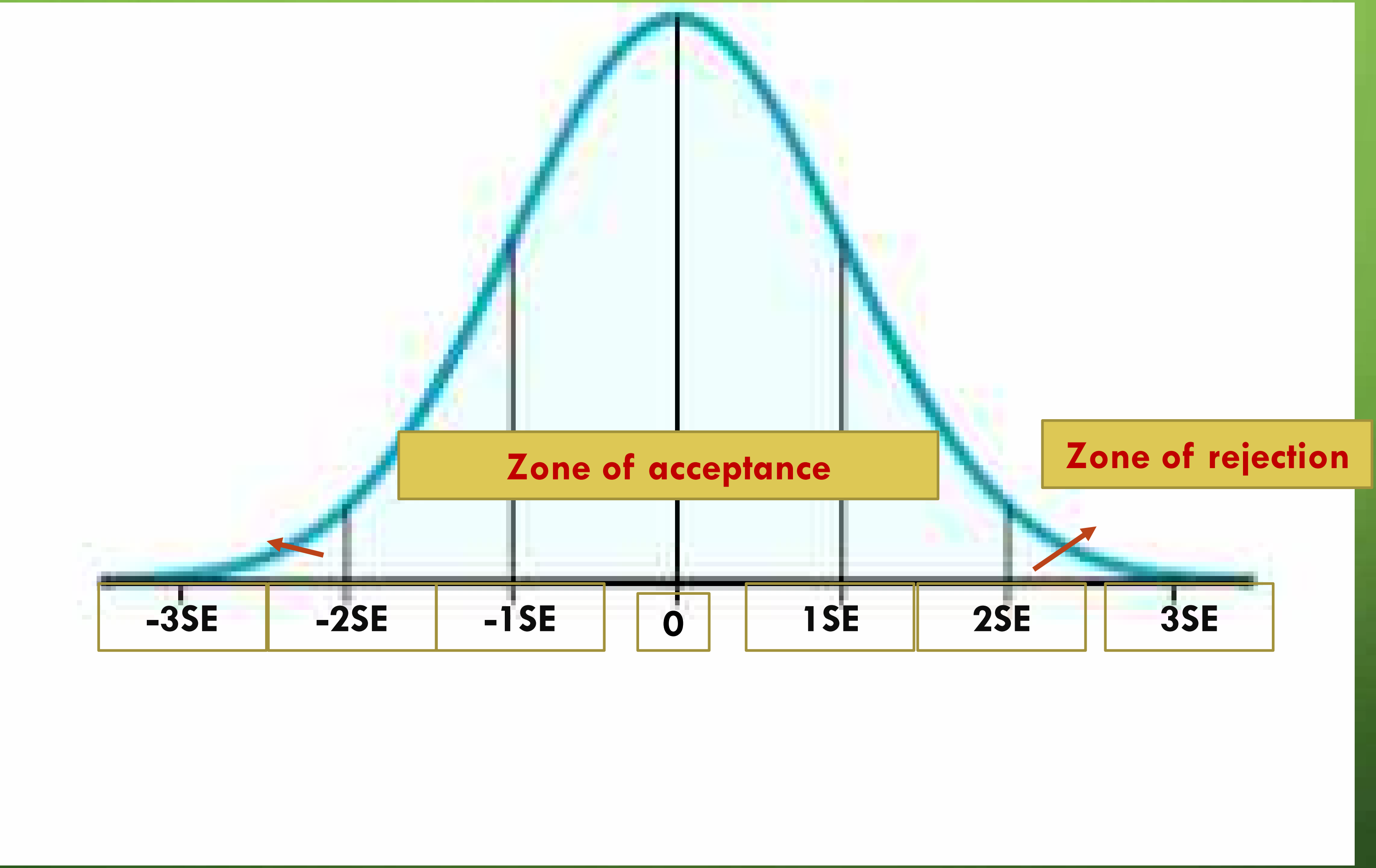
Relative deviate: deviation or distance of an observed value from the mean in terms of standard deviation. (Position of an observed value or observed difference away from the mean in terms of SD.)

- Relative deviate Z= $\frac{x-\bar{x}}{SD}$
- One sample Z= $\frac{\bar{x}-\mu}{SE}$
- Two sample means z= $\frac{\bar{x}_1-\bar{x}_2}{SE\ (\bar{x}_1-\bar{x}_2)}$

STANDARD NORMAL CURVE



STANDARD NORMAL CURVE



P VALUE

- Probability of occurrence of a value/difference by chance.
- Probability of getting a value as extreme as or more extreme than the one observed when the null hypothesis is true.
- Probability is relative frequency of occurence of an event.

- P value is calculated from normal distribution curve.
- Eg. the probability of getting a value above or below mean±2SD is 5%. (95% are within those limits)
- the probability of getting a value above or below mean±3SD is only 1%.

- In the normal distribution curve of sample means,

95% of sample means lie within the mean$\pm$2SE limits. The probability of values falling within these range is 95%. The probability of values being higher or lower than this range is 5%. These limits are called **95% confidence limits.**

- P value to be significant,
- <0.05- difference is statistically significant

Student ‘t’ test

STUDENT'S T TEST

- To compare means of two groups.
- Designed by WS **Gossett** whose pen name was **student.**
- Small samples ($n<30$) do not follow normal distribution as large samples do.
- To test the significance of difference between means of small samples.
- Level of significance or P value is determined by using 't' table.

Criteria for applying t test

- Random samples
- Quantitative data
- Normally distributed variable
- Sample size <30

A. ONE SAMPLE STUDENT T TEST

- Compare the mean of a single group of observations with a specified or reference value.
- eg: comparison of mean dietary intake of pregnant woman with the recommended daily intake.

B. TWO SAMPLE STUDENT T TEST (UNPAIRED & PAIRED)

- **Unpaired t test- Two independent samples**
- $t = \dfrac{\bar{x}_1 - \bar{x}_2}{SE\ (\bar{x}_1 - \bar{x}_2)}$
- SE=c√1/n1+ 1/n2
- df= n1 +n2 -2

B. TWO SAMPLE STUDENT T TEST (UNPAIRED & PAIRED)

Paired t test; observations made on one sample before and after intervention.

- t= $\frac{\bar{x}-0}{SE}$
- SE= SD/√n
- df= n-1

ANALYSIS OF VARIANCE (ANOVA) TEST

- To compare means of more than 2 groups.

Assumptions:

- Samples are random & independent of each other
- More than 2 groups in independent variable
- Groups should have equal variances
- Outcome variable is normally distributed.

ANALYSIS OF VARIANCE (ANOVA) TEST

- F ratio is calculated to find the significance of difference.
- When between group variability is more than the within the group variability, the difference is statistically significant.

EXERCISES