F-test

Variance ratio test

F-test-Variance ratio test

- **The test based on the test statistic which follows F-distribution is called F- test**

Significant difference between variances of two populations based on small samples drawn from those populations

- $$F = \frac{\dfrac{s_1^2 n_1}{n_1 - 1}}{\dfrac{s_2^2 n_2}{n_2 - 1}}$$

- In the F-ratio we always take the larger of the two estimates in the numerator and smaller in the denominator
- The degree of freedom is (n_1-1) and (n_2-1)

Assumptions

- Samples are drawn at random
- Samples are drawn from normal populations
- Population standard deviations are treated as equal but unknown

uses

- Test equality of variances of two populations
- To test equality of means of three or more populations (known as analysis of variance)

HYPOTHESIS TESTING

In this session ....

- What is hypothesis testing?
- Interpreting and selecting significance level
- Type I and Type II errors
- One tailed and two tailed tests
- Hypothesis tests for population mean
- Hypothesis tests for population proportion
- Hypothesis tests for population standard deviation

What is Hypothesis Testing?

Hypothesis testing refers to

1. Making an assumption, called hypothesis, about a population parameter.
2. Collecting sample data.
3. Calculating a sample statistic.
4. Using the sample statistic to evaluate the hypothesis (how likely is it that our hypothesized parameter is correct. To test the validity of our assumption we determine the difference between the hypothesized parameter value and the sample value.)

HYPOTHESIS TESTING

Null hypothesis, H_0

- State the hypothesized value of the parameter before sampling.
- The assumption we wish to test (or the assumption we are trying to reject)
- E.g population mean $\mu = 20$
- There is no difference between coke and diet coke

Alternative hypothesis, H_A

All possible alternatives other than the null hypothesis.

E.g $\mu \neq 20$

$\mu > 20$

$\mu < 20$

There is a difference between coke and diet coke

Null Hypothesis

The null hypothesis H_0 represents a theory that has been put forward either because it is believed to be true or because it is used as a basis for an argument and has not been proven. For example, in a clinical trial of a new drug, the null hypothesis might be that the new drug is no better, on average, than the current drug. We would write

H_0: there is no difference between the two drugs on an average.

Alternative Hypothesis

The alternative hypothesis, H_A, is a statement of what a statistical hypothesis test is set up to establish. For example, in the clinical trial of a new drug, the alternative hypothesis might be that the new drug has a different effect, on average, compared to that of the current drug. We would write

H_A: the two drugs have different effects, on average.

or

H_A: the new drug is better than the current drug, on average.

The result of a hypothesis test:

'Reject H_0 in favour of H_A' OR 'Do not reject H_0'

Selecting and interpreting significance level

1. Deciding on a criterion for accepting or rejecting the null hypothesis.
2. Significance level refers to the percentage of sample means that is outside certain prescribed limits. E.g testing a hypothesis at 5% level of significance means

- that we reject the null hypothesis if it falls in the two regions of area 0.025.
- Do not reject the null hypothesis if it falls within the region of area 0.95.

3. The higher the level of significance, the higher is the probability of rejecting the null hypothesis when it is true. (acceptance region narrows)

Type I and Type II Errors

1. Type I error refers to the situation when we reject the null hypothesis when it is true (H_0 is wrongly rejected).
   e.g H_0: there is no difference between the two drugs on average.
   Type I error will occur if we conclude that the two drugs produce different effects when actually there isn't a difference.
   Prob(Type I error) = significance level = α
2. Type II error refers to the situation when we accept the null hypothesis when it is false.
   H_0: there is no difference between the two drugs on average.
   Type II error will occur if we conclude that the two drugs produce the same effect when actually there is a difference.
   Prob(Type II error) = ß

Type I and Type II Errors – Example

Your null hypothesis is that the battery for a heart pacemaker has an average life of 300 days, with the alternative hypothesis that the average life is more than 300 days. You are the quality control manager for the battery manufacturer.

(a) Would you rather make a Type I error or a Type II error?

(b) Based on your answer to part (a), should you use a high or low significance level?

Type I and Type II Errors – Example

Given H_0 : average life of pacemaker = 300 days, and H_A: Average life of pacemaker > 300 days

(a) It is better to make a Type II error (where H_0 is false i.e average life is actually more than 300 days but we accept H_0 and assume that the average life is equal to 300 days)

(b) As we increase the significance level (α) we increase the chances of making a type I error. Since here it is better to make a type II error we shall choose a low α.

Two Tail Test

Two tailed test will reject the null hypothesis if the sample mean is significantly higher or lower than the hypothesized mean.

Appropriate when $H_0 : \mu = \mu_0$ and $H_A: \mu \neq \mu_0$

e.g The manufacturer of light bulbs wants to produce light bulbs with a mean life of 1000 hours. If the lifetime is shorter he will lose customers to the competition and if it is longer then he will incur a high cost of production. He does not want to deviate significantly from 1000 hours in either direction. Thus he selects the hypotheses as

$H_0 : \mu = 1000$ hours and $H_A: \mu \neq 1000$ hours

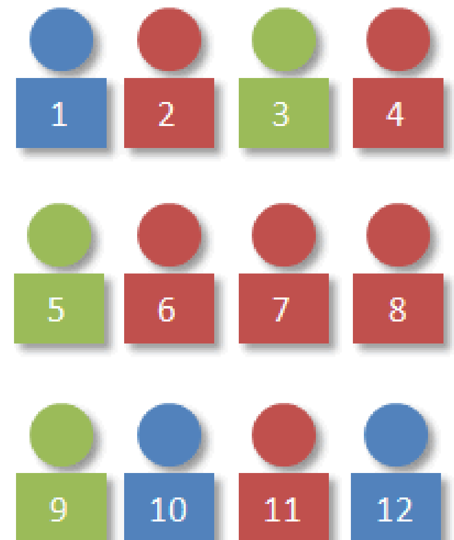
and uses a two tail test.

Sampling Plan

Things to Discuss

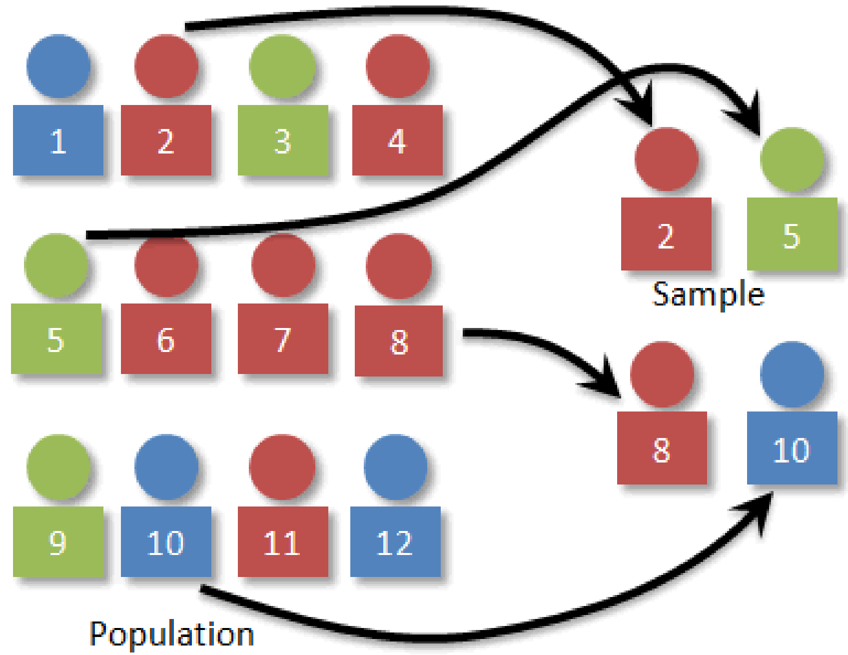
- Key Terminologies
- Sampling Design Process
- Sampling Classification (Scheme)
- Different Sampling Methods
- Sample size Calculation

Population



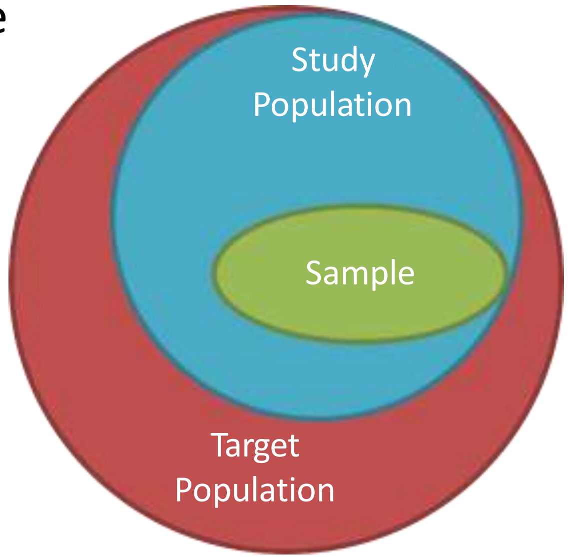
Population

Sample



Key Terminologies

- Population:
  - a set which includes all measurements of interest to the researcher (The collection of all responses, measurements, or counts that are of interest)
- Sample:
  - A subset of the population



Key Terminologies

- Study(Sampling) Population
  - The population to be studied/ to which the investigator wants to generalize his results
- Sampling Unit
  - smallest unit from which sample can be selected
- Sampling frame
  - List of all the sampling units from which sample is drawn
- Sampling scheme
  - Method of selecting sampling units from sampling frame
- Sampling fraction
  - Ratio between sample size and population size

Sampling Process

1. Define the population
2. Determine the sampling frame
3. Select Sampling Techniques
4. Determine the Sample size
5. Execute Sampling Process

Influencing Factors

- Factors that influence sample design
  - Research Type
  - Element Type
  - Population size
  - Participation (response)
  - Available Resources
  - Constraints/ limitations

- When might you sample the entire population?
  - When your population is very small
  - When you have extensive resources
  - When you don’t expect a very high response

Sampling Classifications

Probability Sampling

- Simple random sample
- Systematic sample
- Stratified sample
- Cluster sample
- Multistage sample

Each member of the population has a known non-zero probability of being selected, and this probability can be accurately determined.

Non-Probability Sampling

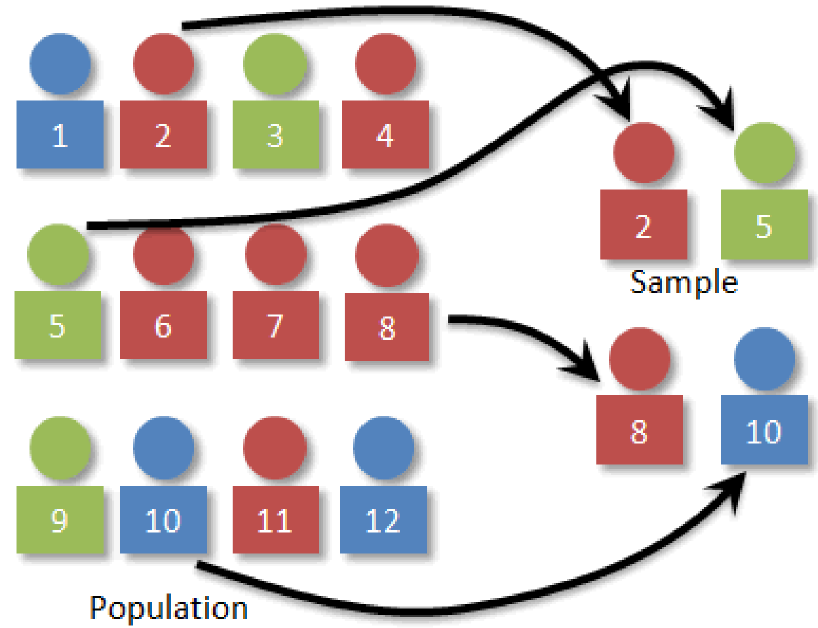
- Convenience sample
- Judgmental/ Purposive sample
- Quota sample
- Snowball sample

Members are selected from the population in some nonrandom manner or where the probability of selection can't be accurately determined

PROBABILITY SAMPLING

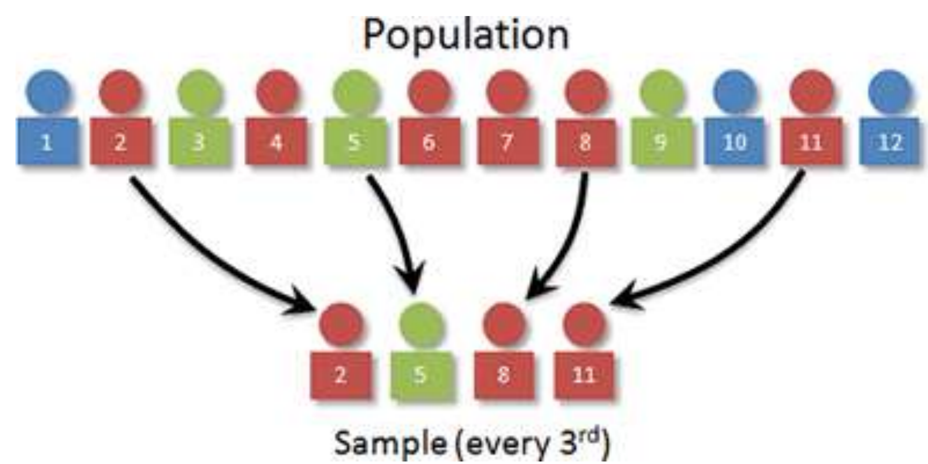
Simple Random Sampling

- Each element in the population has a known and equal probability of selection.



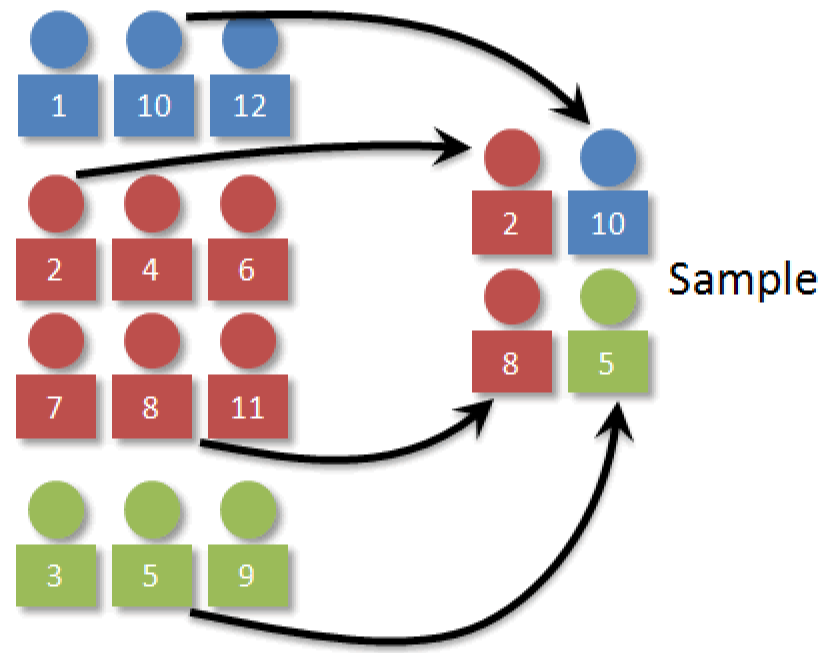
Systematic Sampling

- Arranging the target population according to some ordering scheme and then selecting elements at regular intervals through that ordered list.



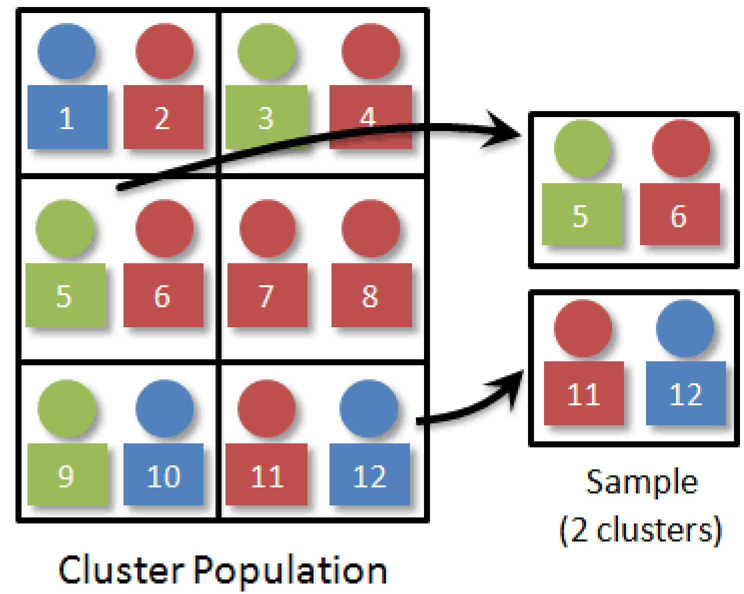
Stratified Sampling

- Population with some distinct categories can be organized into separate "stratum" which can be sampled as an independent sub-population, out of which individual elements can be randomly selected.

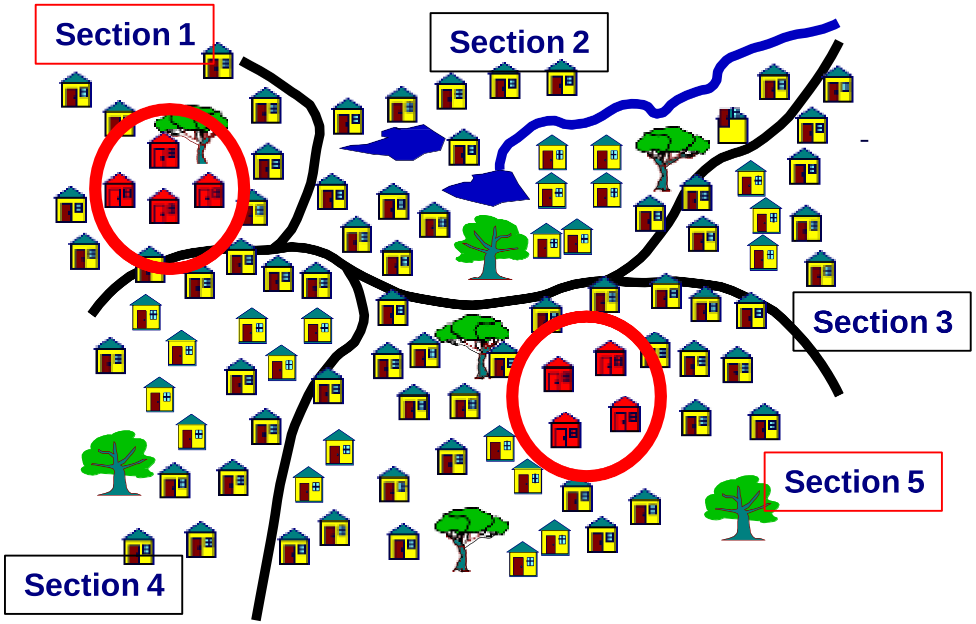


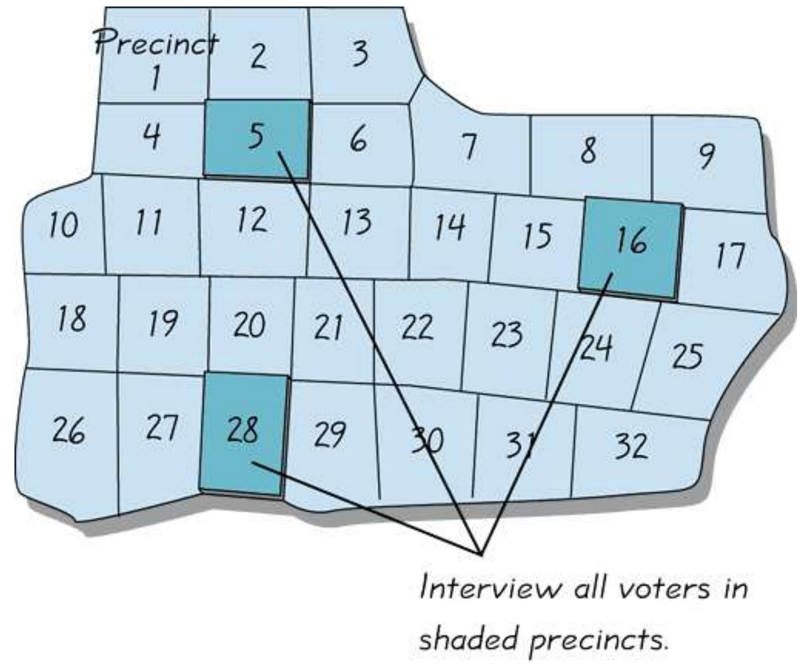
Cluster Sampling

- Divide the population into groups (called clusters), randomly select some of the groups, and then collect data from ALL members of the selected groups

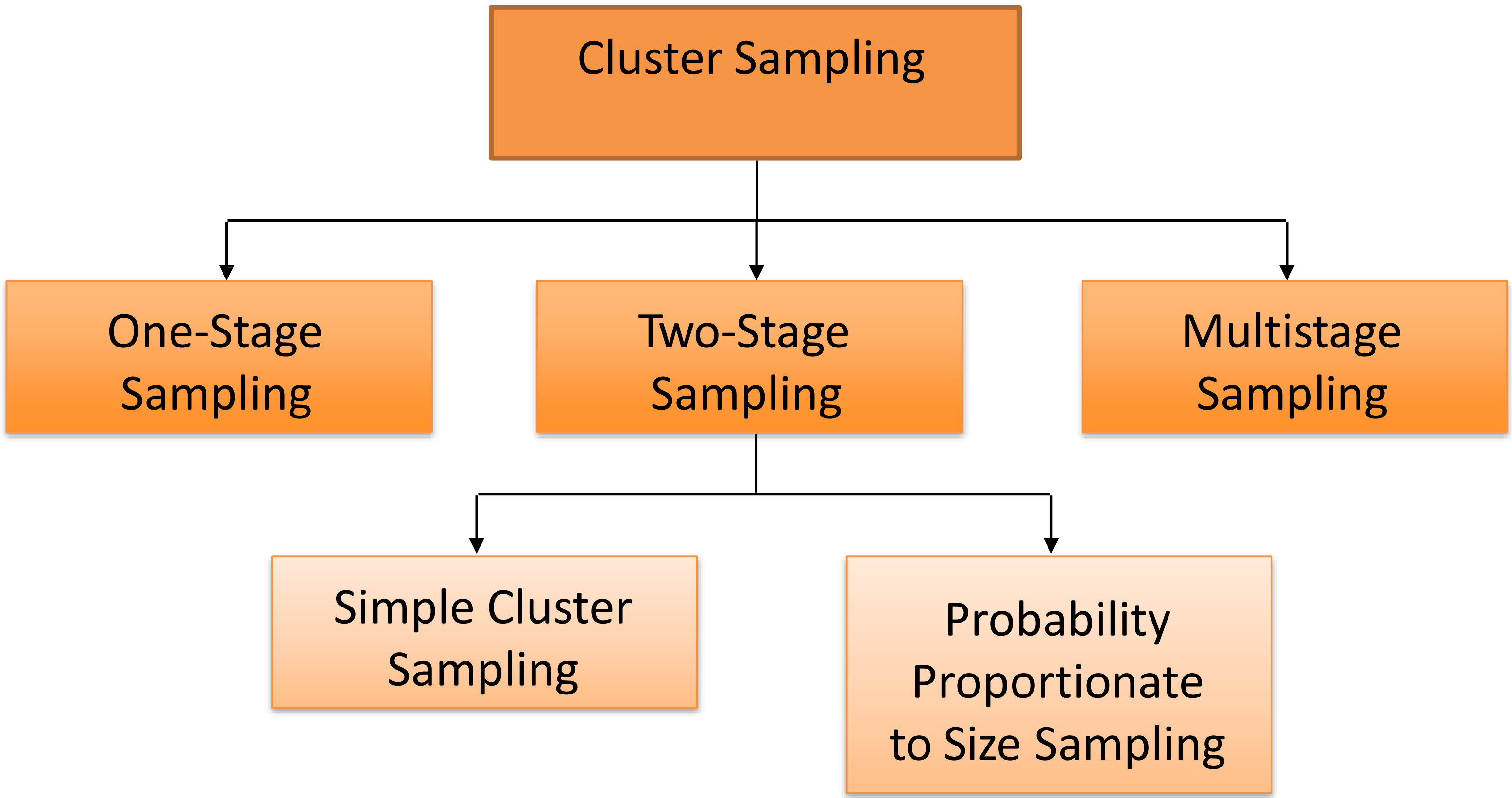


Cluster Sampling





Multistage Sampling



NON-PROBABILITY SAMPLING

Convenience Sampling

- Convenience sampling attempts to obtain a sample of convenient elements. Often, respondents are selected because they happen to be in the right place at the right time.
  - use of students, and members of social organizations
  - mall intercept interviews without qualifying the respondents
  - department stores using charge account lists
  - "people on the street" interviews

Purposive (Judgmental) Sampling

- Judgmental sampling is a form of convenience sampling in which the population elements are selected based on the judgment of the researcher.
  - test markets
  - purchase engineers selected in industrial marketing research
  - bellwether precincts selected in voting behavior research
  - expert witnesses used in court

Quota Sampling

- Quota sampling may be viewed as two-stage restricted judgmental sampling.
  - The first stage consists of developing control categories, or quotas, of population elements.
  - In the second stage, sample elements are selected based on convenience or judgment.

Control Characteristic	Population composition Percentage	Sample composition Percentage	Number
Sex			
Male	48	48	480
Female	52	52	520
	100	100	1000

Snowball Sampling

- In snowball sampling, an initial group of respondents is selected, usually at random.
  - After being interviewed, these respondents are asked to identify others who belong to the target population of interest.
  - Subsequent respondents are selected based on the referrals.

Strengths and Weaknesses of Basic Sampling Techniques

Technique	Strengths	Weaknesses
Nonprobability Sampling		
Convenience sampling	Least expensive, least time-consuming, most convenient	Selection bias, sample not representative, not recommended for descriptive or causal research
Judgmental sampling	Low cost, convenient, not time-consuming	Does not allow generalization, subjective
Quota sampling	Sample can be controlled for certain characteristics	Selection bias, no assurance of representativeness
Snowball sampling	Can estimate rare characteristics	Time-consuming
Probability sampling		
Simple random sampling (SRS)	Easily understood, results projectable	Difficult to construct sampling frame, expensive, lower precision, no assurance of representativeness.
Systematic sampling	Can increase representativeness, easier to implement than SRS, sampling frame not necessary	Can decrease representativeness
Stratified sampling	Include all important subpopulations, precision	Difficult to select relevant stratification variables, not feasible to stratify on many variables, expensive
Cluster sampling	Easy to implement, cost effective	Imprecise, difficult to compute and interpret results

Errors in sample

- Systematic error (or bias)
  - Inaccurate response (information bias)
  - Selection bias

- Sampling error (random error)

“t” Test

Overview

- Background
- Different versions of t-test
- Main usage of t-test
- t-test v/s z-test
- Assumptions of t-test
- Examples

Background

- Introduced in 1908 by William Sealy Gosset.
- Gosset published his mathematical work under the pseudonym “Student”.
- Definition of t test: “ It’s a method of testing hypothesis about the mean of small sample drawn from a normally distributed population when the standard deviation for the sample is unknown.”

Assumptions of t-Test

- Dependent variables are interval or ratio.
- The population from which samples are drawn is normally distributed.
- Samples are randomly selected.
- The groups have equal variance (Homogeneity of variance).
- The t-statistic is robust (it is reasonably reliable even if assumptions are not fully met.)

Applications of t test

- The calculation of a confidence interval for a sample mean.
- To test whether a sample mean is different from a hypothesized value.
- To compare mean of two samples.
- To compare two sample means by group.

Types of "t" test

- Single sample t test – we have only 1 group; want to test against a hypothetical mean.

- Independent samples t test – we have 2 means, 2 groups; no relation between groups, Eg: When we want to compare the mean of T/T group with Placebo group.

- Paired t test – It consists of samples of matched pairs of similar units or one group of units tested twice. Eg: Difference of mean pre & post drug intervention.

One Sample t-test

- It is used in measuring whether a sample value significantly differs from a hypothesized value.

- For example, a research scholar might hypothesize that on an average it takes 3 minutes for people to drink a standard cup of coffee.

- He conducts an experiment and measures how long it takes his subjects to drink a standard cup of coffee.

- The one sample t-test measures whether the mean amount of time it took the experimental group to complete the task varies significantly from the hypothesized 3 minutes value.

Equation for a one-sample t-test

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

where

t = the t statistic

$\bar{x}$ = the mean of the sample

μ = the comparison mean

s = the sample standard deviation

n = the sample size

Example

- 10 individuals had taken an exam and we want to test whether their scores, all together, are significantly different from the score of 100.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{107.8 - 100}{5.35/\sqrt{10}} = 4.61$$

•We need to calculate the degrees of freedom.
•Here, the degrees of freedom is simply the sample size minus one.
•Therefore, Degrees of freedom = n – 1 = 10 – 1 = 9
•Now, we will refer to a t table to determine the critical t value for 9 degrees of freedom at the .05 level of significance.
•Looking at a t table, this value is 2.26 .
•Since our calculated t value of 4.61 is greater than the critical t value of 2.26, we can say that the scores of our sample of 10 individuals differ significantly from the score of 100.

T table

df	.05	.01	.001
1	12.706	63.657	636.619
2	4.303	9.925	31.598
3	3.182	5.841	12.924
4	2.776	4.604	8.610
5	2.571	4.032	6.869
6	2.447	3.707	5.959
7	2.365	3.499	5.408
8	2.306	3.355	5.041
9	2.262	3.250	4.781
10	2.228	3.169	4.587

Independent t test

- The independent sample t-test consists of tests that compare mean value(s) of continuous-level (interval or ratio data), in a normally distributed data.
- The independent sample t-test compares two means.
- The independent samples t-test is also called unpaired t-test/ two sample t test.
- It is the t-test to be used when two separate independent and identically distributed variables are measured.
- Eg: 1. Comparision of quality of life improved for patients who took drug Valporate as opposed to patients who took drug Levetiracetam in myoclonic seizures.
- 2.Comparasion of mean cholesterol levels in treatment group with placebo group after administration of test drug.

Assumptions

- A random sample of each population is used.
- The random samples are each made up of independent observation.
- Each sample is independent of one another.
- The population distribution of each population must be nearly normal, or the size of the sample is large.

Independent t test

To test the null hypothesis that the two population means, μ1 and μ2, are equal:

- 1. Calculate the difference between the two sample means, $\bar{x}_1 - \bar{x}_2$.
- 2. Calculate the pooled standard deviation: sp
- 3. Calculate the standard error of the difference between the means:
- 4. Calculate the T-statistic, which is given by $T = \bar{x}_1 - \bar{x}_2 / SE(\bar{x}_1 - \bar{x}_2)$
- This statistic follows a t-distribution with $n1 + n2 - 2$ degrees of freedom.
- 5. Use tables of the t-distribution to compare your value for T to the $t_{n1+n2-2}$ distribution. This will give the p-value for the unpaired t-test.

Equation for the independent samples t-test

•The independent-Samples t-test procedure compares means for two groups of cases.

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\left[\frac{SS_1 + SS_2}{n_1 + n_2 - 2}\right]\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

Here,

$\overline{X}_1$ and $\overline{X}_2$ are the means of the two different groups

$n_1 = n$ of Group 1

$n_2 = n$ of Group 2

SS = sum of squares

Example

- Suppose we have to compare the mean value of two groups, one with 7 subjects and the other with 5 subjects .
- These were their scores:

Case	Group 1	Group 2
1	78	87
2	82	92
3	87	86
4	65	95
5	75	73
6	82	
7	71	

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{SS_1 + SS_2}{n_1 + n_2 - 2}\right]\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} = \frac{77.14 - 86.60}{\sqrt{\left[\frac{334.86 + 285.20}{7 + 5 - 2}\right]\left[\frac{1}{7} + \frac{1}{5}\right]}}$$

$$= \frac{-9.46}{\sqrt{\left(\frac{620.06}{10}\right)\left(\frac{12}{35}\right)}} = \frac{-9.46}{\sqrt{21.26}} = -0.44$$

For an independent or between subjects' t test: df = n1+ n2 - 2

- Now, take the absolute value of this, which is 0.44.
- Now, for the .05 probability level with 10 degrees of freedom, we see from the table that the critical t score is 2.228 for a two-tailed test.
- Since the calculated t score is lower than the critical t score, the results are not significant at the .05 probability level.

T table

df	.05	.01	.001
1	12.706	63.657	636.619
2	4.303	9.925	31.598
3	3.182	5.841	12.924
4	2.776	4.604	8.610
5	2.571	4.032	6.869
6	2.447	3.707	5.959
7	2.365	3.499	5.408
8	2.306	3.355	5.041
9	2.262	3.250	4.781
10	2.228	3.169	4.587

Paired t test

- A paired t-test is used to compare two population means where you have two samples in which observations in one sample can be paired with observations in the other sample.
- A comparison of two different methods of measurement or two different treatments where the measurements/treatments are applied to the same subjects.
- Eg: 1.pre-test/post-test samples in which a factor is measured before and after an intervention,
- 2. Cross-over trials in which individuals are randomized to two treatments and then the same individuals are crossed-over to the alternative treatment,
- 3. Matched samples, in which individuals are matched on personal characteristics such as age and sex,