04-03-2021

Unit-III Sampling And Sampling Distribution

Population:- The set of all objects under the consideration is called population.

Sample:- A small portion of a population is called a sample.

Finite Population:- A population with finite no. of elements is called finite population.

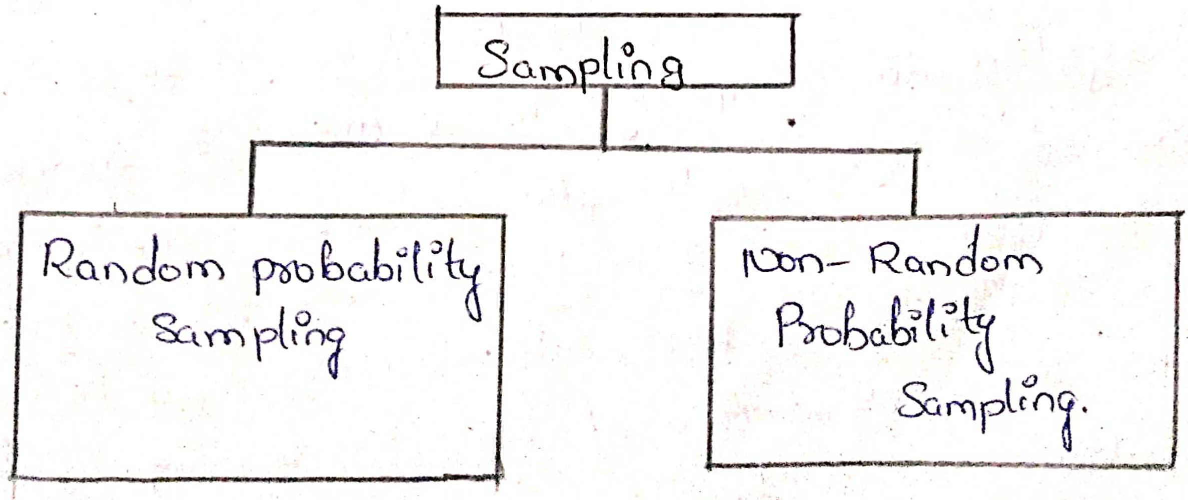
Infinite Population:- A population whose size is very large is called infinite population.

A population whose size is not given is considered as infinite population.

Large Sample:- A ~~la~~ sample with size more than (or) equal to 30 that is $n \geq 30$ is called a large sample.

Small Sample:- A sample with size less than 30 i.e., $n < 30$ is called small sample.

Sample survey:- The process of collecting information from the sample is called sample survey.



Random Sampling / Random Probability Sampling:-

A method of selecting a sample from population in which all the items in the population have a equal chance of being choosen in the given sample.

This implies that the selection of sample items is independent of person who is making the investigation.

There are 4 methods of Random Probability Sampling:-

1) Simple random Sampling
2) Systematic random Sampling
3) Stratified random Sampling
4) cluster random sampling.

1) Simple Random Sampling:-

In this method each individual observation/item of the population as an equal chance of being picked in the sample. This method is used for when the population is small and the inform is readyly avialable such as, no. of ~~cos~~ customers coming to a restuarent and, no. of students in a particular class and so on --

2) Systamatic Sampling:- These means the Sample is obtained in some systamatic manner. usally by taking items at particular given intervals regularly in this case all the items of population are arranged in some order. This method is used for when a complete list of Population is avialable.

Ex:- Consider there 1000 persons from whom we have to choose 10 persons for the study of any given Sample then we number them from 1 to 1000 and make them as 10 intervals and 1st we choose a person from the 1st interval 2nd person is choosen from the 2nd interval and so on ---.

As we have numbered the persons all the persons are systematically choosen. This kind of method is used when the population is large.

Stratified Sampling:- It is generally used when the Population is not small in this method the population is first sub-divided into several small groups and these small groups are called stratal according to some relevant characteristics so that each group is more or less will be same in size now the sample is selected by drawing individuals from each stratal at a random space. In each of its proportion size. This is called stratified sample.

Clustar Sample:- A clustar sample randomly selected group this method is useful when the population quietly ~~dispersed~~ dispersed and consists of many natural group such as factories, villages, e.t.c --

Note:- The statistical inference is the process by which the inferences about a population are made from the information of a sample.

Parameter:- The characteristic of a population is called Parameter. (44)

Statestic:- The characteristic of a ~~por~~ Sample is called statestic.

Notations:-

	Population	Sample
Size	N	n
Mean	μ	$\bar{x}$
Proportion	P	p
S.D (standard deviation)	σ	s

Sampling distribution of mean:-

The probability distribution of all the possible sample means from a population is called sampling distribution of mean.

Sampling distribution of Proportion:-

The probability distribution of all the possible Sample proportions from a population is called Sampling distribution of proportion.

Standard Error of mean $\bar{x}$:- $(\sigma_{\bar{x}})$

The standard deviation of sampling distribution of the mean is called standard error of mean.

Formula:- $\boxed{\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}}$ (when population is infinite)

2) $\boxed{\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}}$ (when population is finite)

Standard error of proportion:-

The standard deviation of sampling distribution of a given proportion is called standard error of proportion.

Formula:- $\sigma_p = \sqrt{\frac{pq}{n}}$

Standard Error:-

While calculating the standard deviation of the sample distribution of a statestic is called standard. Error.

Sampling Fraction:-

The fraction/Proportion of the population contained in a sample is called sampling Fraction.

Finite Population:- Its formula is given by $\sqrt{\frac{N-n}{N-1}}$ is called finite population multiplier, where,

N = Population size; n = sample size.

Central Limit Theorem:-

Let n be size of sample taken from a population having mean(μ) then sampling distribution of mean's approaches to normality and mean of the sampling distribution of means is equal to the population mean when $n \geq 30$.

Let 'n' be the size of sample taken from a population having proportion "π" then sampling distribution of proportions approaches to normality and mean of the sampling distribution of their proportions is equal to population proportion when $n \geq 30$.

1) From a population of 125 items, with the mean of 105 and standard deviation of 17 among them 64 items were choosen the

(46)

a) what is the standard error of mean :-

b) what is the probability of $P(107.5 < \bar{x} < 109)$

Sol:- Given that

Sample size $n = 64$

$N = 125$

(mean) $\mu = 105$

(standard deviation) $\sigma = 17$

a) we know that

standard error of mean formula when there is finite

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{17}{\sqrt{64}} \sqrt{\frac{125-64}{125-1}}$$

$$= 2.125\sqrt{0.4919}$$

$$\sigma_{\bar{x}} = 1.4903$$

b) $Z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}}$

$= \frac{107.5 - 105}{1.4903}$

$= 1.68$

$Z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}}$

$= \frac{109-105}{1.4903}$

$= 2.68$

$P(1.68 < Z < 2.68)$

$= P(0 < Z < 2.68) - P(0 < Z < 1.68)$

$= 0.4963 - 0.4535$

$= 0.0428$

2) From a population of 75 items, with the mean of 364 and ~~standard~~ variance of 18; 32 items were randomly selected without replacement (47)

a) what is the standard error of mean.

b) find probability of $(363 \le \bar{x} \le 366)$

Sol:- given N(population) = 75

$\mu = 364$

(variance) $\sigma^2 = 18$ $\quad n = 32$

$\sigma = \sqrt{18}$

$= 4.24$

a) standard error mean value

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

$$= \frac{4.24}{\sqrt{32}} \sqrt{\frac{75-32}{75-1}}$$

$$= \frac{4.24}{5.65} \sqrt{\frac{43}{74}}$$

$$= 0.7504(0.7622)$$

$$= 0.571$$

b) $Z = \dfrac{\bar{x}-\mu}{\sigma_{\bar{x}}}$

$= \dfrac{363-364}{0.571}$

$= -1.75$

$Z = \dfrac{\bar{x}-\mu}{\sigma_{\bar{x}}}$

$= \dfrac{366-364}{0.571}$

$= 3.50$

$$P(-1.75 < \bar{x} < 3.50) = P(0 < \bar{x} < 1.75) + P(0 < \bar{x} < 3.50)$$

$$= 0.4599 + 0.5$$

$$= 0.9599$$

3) Given a population size $N = 80$, with mean$(\mu) = 22$ and standard deviation of $\sigma = 3.2$ what is the probability that a sample of $n = 25$ will have a mean between 21 & 23.5

Sol:- Given that :-

(48)

$N = 80$; $n = 25$; $\mu = 22$ & $\sigma = 3.2$

to find $P(21 \leq \bar{x} \leq 23.5)$

a) standard error of mean :- $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$

$$\frac{3.2}{\sqrt{25}} \sqrt{\frac{80-25}{80-1}}$$

$$= 0.64 \sqrt{\frac{55}{79}} \Rightarrow 0.64 (0.834)$$

$$= 0.5312$$

b) $P(21 \leq \bar{x} \leq 23.5)$

$$z = \frac{\bar{x} - \mu}{\sigma_x} \qquad z = \frac{\bar{x} - \mu}{\sigma_x}$$

$$\frac{21 - 22}{0.53} \qquad \frac{23.5 - 22}{0.53}$$

$$= -1.88 \qquad = 2.830$$

$P(-1.88 \leq \bar{x} \leq 2.830) = P(0 \leq z \leq 1.88) + P(0 \leq z \leq 2.83)$

$= 0.4699 + 0.4977$

$= 0.9676$

4) In a sample of 25 observations from a normal distribution with mean$(\mu) = 98.6$, and a standard deviation (σ) of $\sigma = 17.2$. what is the probability of $P(92 < \bar{x} < 102)$

Sol:- $n = 25$, $\mu = 98.6$, $\sigma_x = 17.2$

find $p(92 < \bar{x} < 102)$

a) standard Error of mean when N is infinity

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{17.2}{\sqrt{25}} = \frac{17.2}{5} = 3.44$$

b) $p(92 < \bar{x} < 102)$

$$Z_1 = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{92 - 98.6}{3.44} = -1.941$$

$$Z_2 = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{102 - 98.6}{3.44} = 0.98$$

$$P(-1.941 < x < 0.98) = (0 < x < 0.98) + (0 < x < 1.94)$$

$$= 0.3365 + 0.4738$$

$$= 0.8103$$

5) From a population of 120 items with a mean of 100 and a S.D = 17, 60 items are choosen then,

a) what is the standard error of mean

b) what is $p(103 < \bar{x} < 104)$

Sol:- given $N = 120$, $n = 60$, $\mu = 100$ and $\sigma = 17$

a) $$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{17}{\sqrt{60}}\sqrt{\frac{120-60}{120-1}}$$

$$= 2.19\sqrt{\frac{60}{119}}$$

$$= 2.19(0.71)$$

$$= 1.55$$

b) $p(103 < \bar{x}\ 104)$

$Z_1 = \frac{103-100}{1.55} = \frac{3}{1.55}$, $Z_2 = \frac{104-100}{1.55} = \frac{4}{1.55}$

$= 1.93$ $= 2.58$

$p(1.93 < z < 2.58) = p(0 < z < 2.58) - p(0 < z < 1.93)$

$= 0.4951 - 0.4732$

$= 0.0219$

5) 70 data clerks at the department of motor vechicles make an average of 18 errors per day Normally distributed the standard deviation $=4$, a field Auditor can check the work 15 clerk's per day what is the probability that a average No. of errors in a group 15 clerks checked on one day is

a) fewer than 15.5 i.e., $P(\bar{x} \leq 15.5)$

b) Greater than 20 i.e., $p(\bar{x} > 20)$

Sol:- $N = 70, n = 15, \sigma = 4, \mu = 18$

standard error of mean.

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{4}{\sqrt{15}}\sqrt{\frac{70-15}{70-1}}$

$= 1.03\sqrt{\frac{55}{69}} = 1.03(0.8)$

$= 0.824$

a) $p(\bar{x} \leq 15.5)$

$z = \frac{15.5-18}{0.824} = \frac{-2.5}{0.824}$

$= -3.03$

$= 0.5 - p(0 \leq z \leq -3.03)$

$= 0.5 + p(0 \leq z \leq 3.03)$

$= 0.5 + 0.4988$

$= 0.9988$

b) $p(\bar{x} > 20)$

$z = \frac{20-18}{0.824} = \frac{2}{0.824}$

$= 2.42$

$= 0.5 - p(z > 2.42)$

$= 0.5 - P(0 < z < 2.42)$

$= 0.5 - 0.4922$

$= 0.0078$

18-03-2021

Estimations:-

1) Point estimate:- A single number used to estimate an unknown population parameter, is called point estimate

Interval Estimate:- A range of values used to estimate an unknown population parameter is called

Interval estimate.

Confidence level:- The probability that statistician associates with an interval estimate of a population parameter indicating how confident they are the interval estimate will include the population parameter.

Confidence interval:-

A range of values that has some probabilities of including the true population parameter is called

confidence interval.

confidence intervals for population mean is given by.

$$\boxed{\bar{x} \pm z\,\sigma_{\bar{x}}}$$

where, $\bar{x}$ = sample mean

$\sigma_{\bar{x}}$ = S.E of mean

z = normal distribution values for the given confidence

Confidence Intervals:-

C.L	90%	95%	99%
Z	1.645	1.96	2.58

(Or:ca)

Ex:- 1) A sample of 900 members is selected from a normal distribution the mean and s.d of a sample are 3.4cm and 2.61cm obtain 99% confidence interval by the population mean.

52

Sol:- Given $n = 900$, $x = 3.4cm$, $\sigma_{\bar{x}} = 2.61 cm$

Confidence level (C.L) = 99%

Confidence interval of population mean is $\bar{x} \pm z\sigma_{\bar{x}}$

At 99% of C.L $Z = 2.58$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2.61}{\sqrt{900}} = \frac{2.61}{30} = 0.087$$

$\bar{x} \pm z\sigma_{\bar{x}}$

$= 3.4 \pm 2.58(0.087) = 3.4 \pm 0.2245$

$(3.4 - 0.2245)$ to $(3.4 + 0.2245)$

$= 3.175$ to 3.6245

C.I of given mean is $(3.175$ to $3.6245)$

2) For a population known you have s.d = 1.4 and sample of 60 individuals taken and the mean for this sample is 6.2 find 90%, 95%, 99% of C.I's

given:- $n = 60$; $\bar{x} = 6.2$, $\sigma = 1.4$

at 90% of C.L $z = 1.645$

$$\sigma_{\bar{x}} = \frac{1.4}{\sqrt{60}} = 0.181$$

$\bar{x} \pm z\sigma_{\bar{x}}$

$6.2 \pm 1.645(0.181)$

6.2 ± 0.298

$6.2 - 0.298$ to $6.2 + 0.298$

$\therefore$ C.I = (5.902 to 6.498)

At 95% of C.L $Z = 1.96$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = 0.181$$

$$\bar{x} \pm z\,\sigma_{\bar{x}} = 6.2 \pm 1.96(0.181)$$

$$6.2 \pm 0.354$$

$$6.2 - 0.354 \text{ to } 6.2 + 0.354$$

$$5.846 \text{ to } 6.554$$

At 99% of C.L $Z = 2.58$

$$\bar{x} \pm z\,\sigma_{\bar{x}} = 6.2 \pm 2.58(0.181)$$

$$6.2 \pm 0.466$$

$$6.2 - 0.466 \text{ to } 6.2 + 0.466$$

$$(5.734) \text{ to } (6.666) = (5.734, 6.666)$$

3) From a population of 540 a sample of 60 individuals are taken from this sample the mean is found to be 6.2 and the σ is 1.386

1) find the estimated standard error of mean.

2) Construct the 96% C.I of the mean.

Sol:- Given $N = 540$ $n = 60$; $\bar{x} = 6.2$; $\sigma = 1.386$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{1.386}{\sqrt{60}} \sqrt{\frac{540-60}{540-1}}$$

$$= 0.1789 \sqrt{\frac{480}{539}}$$

$$= 0.1789(0.943)$$

$$= 0.168$$

At 96% of C.L $z = 2.06$ $\left[\because \frac{96}{100} = \frac{0.96}{2} = 0.48 = 2.06 \text{ (by N.D table)}\right]$

$6.2 \pm 2.06(0.168)$

6.2 ± 0.3460

$6.2 - 0.3460$ to $6.2 + 0.3460$

$= 5.854$ to $6.546 = (5.854, 6.546)$

4) From a population with $\sigma = 16.5$ a sample of 32 items resulted in 34.8 as a estimated of the mean compute an interval estimate that should include the population mean with 99.7%.

Sol:- given $\sigma = 16.5$, $n = 32$, $\bar{x} = 34.8$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{16.5}{\sqrt{32}} = 2.920$$

given C.I = 99.7%

$$\therefore \frac{99.7}{100} \Rightarrow 0.997$$

$$\frac{0.997}{2} = 0.4985$$

$= 2.96$ ($\because$ N.D tables)

$\therefore Z = 2.96$ ($\because$ by normal distribution table)

$x \pm Z\sigma_x \Rightarrow 34.8 \pm (2.96)(2.920)$

34.8 ± 8.6432

$34.8 - 8.6432$ to $34.8 + 8.6432$

$= 26.1568$ to 43.4432

$(26.1568, 43.4432)$

5) For a population with the known variance of 185, a sample of 64 individuals leads to 217 as an estimate of the mean. (55)

a) Find the standard error of the mean.

2) Establish an interval estimate that should include population mean 68.3% of the time was used.

Sol:- given $\sigma^2 = 185$, $n = 64$, $\bar{x} = 217$

$$\sigma = \sqrt{185} = 13.601$$

a) $$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{13.601}{\sqrt{64}} = \frac{13.601}{8} = 1.70$$

b) C.I of 68.3% is

$$Z = \frac{68.3}{100(2)} = \frac{0.683}{2} = 0.3415$$

$= Z = 1.0$ ($\because$ from normal distribution table)

$$\bar{x} \pm Z\,\sigma_{\bar{x}}$$

$= 217 \pm 1.0(1.70)$

$= 217 \pm 1.7$

$= 217 - 1.7$ to $217 + 1.7$

$= 215.3$ to 218.7

$= (215.3, 218.7)$

6) On collecting a sample of 250 from a population with standard deviation of 13.7. The mean is found to be 112.4

a) find 95% 99% C.I for mean.

Sol:- Given, $n=250$, $\bar{x}=112.4$, $\sigma_x=13.7$

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{13.7}{\sqrt{250}} = \frac{13.7}{15.811}$$

$$= 0.866$$

C.I of 95% is 1.96

$\bar{x} \pm z(\sigma_{\bar{x}})$

$= 112.4 \pm (1.96)(0.866)$

112.4 ± 1.697

$(112.4 - 1.697)$ to $(112.4 + 1.697)$

110.703 to 114.097

$(110.703, 114.097)$

C.I of 99% is 2.58

$\bar{x} \pm z \times \sigma_{\bar{x}}$

$112.4 \pm (2.58)(0.866)$

$= 112.4 \pm 2.2342$

$112.4 \pm 2.2342 \rightarrow (112.4 - 2.2342)$ to $(112.4 + 2.2342)$

$= 110.165$ to 114.634

$(110.165, 114.634)$

Standard error of proportion:-

Whenever a sample proportion is given we use the formulae -

$$\boxed{\sigma_p = \sqrt{\frac{pq}{n}}}$$

where, p = Sample proportion
q = 1 − p
n = samesize

Confidence interval for population proportion is,

$$\boxed{p \pm z\sigma_p}$$

where, p = Sample proportion
σ_p = S.E of proportion
z = Normal distribution value for given confidence level.

C.L	90%	95%	99%
z	1.645	1.96	2.58

Ex:-1)
A random sample of 700 units from a large consignment showed 200 for damaged.
i) find 95% confidence limits for the proportion of damaged units in the consignment.
ii) 99% confidence limits for proportion of damaged units in the consignment.

Sol:- Given, n = 700 units in which. 200 units are damaged.
Proportion of damaged units $= \frac{200}{700} = \frac{2}{7} = 0.2857$

i) $P = 0.2857$

$q = 1 - P = 1 - 0.2857$

$= 0.7143$

$\sigma_P = \sqrt{\frac{Pq}{n}}$

$= \sqrt{\frac{(0.2857)(0.7143)}{700}}$

$= \sqrt{\frac{0.204}{700}} = 0.017$

C.L of 99% is 2.58

$P \pm z\,\sigma_p$

$= 0.2857 \pm (2.58)(0.017)$

$= 0.2857 \pm 0.043$

$(0.2857 - 0.043)$ to $(0.2857 + 0.043)$

$= 0.2427$ to 0.3287

$(0.2427, 0.3287)$

ii) C.L at 95% is 1.96

$P \pm z\sigma_p$

$0.2857 \pm (1.96)(0.017)$

0.2857 ± 0.033

$(0.2857 - 0.033)$ to $(0.2857 + 0.33)$

(0.2527) to (0.3187)

$(0, 2527, 0.3187)$