Centeral Limit Theorm - Given a sufficient large Sample size, the Sampling distribution of the mean for a variable will approximately a normal distribution regardless of the variable distribution in the population

Histogram of Average:

Now you make a Histogram or Chart showing the average of these Sample groups. what you will find that according to the Central ~~Tendency~~ Limit Theorm. regardless of the Shape of the original height distribution. this Sample average will start to look like a normal curve.

Normal Distribution Emerges:-

AS you increase the Size of your Sample groups, the distribution of the Sample average become even more like a perfect normal distribution.

Basically

(Please Explain purpose of meeting)

Estimate population using a Z statistic involve. Calculate a confidence interval.

A — Total Cost - Total estimate of the average income for the entire town

- Think of this as the total amount of money earned by everyone in the town

Fixed Cost - This is the initial estimate of average income based on Sample people from town.

(****)

In practice - when we take a sample from our employees every time we do this, we are estimating the mean or average of the entire ~~entire~~ population

4. Draw the Conclusion - Based on the Candies in the bags you checked, you make clusion about the flavours in the entire box

So, Cluster Sampling you are not checking every indiviseual in your population, you are checking a few group and assuming they represent the entire population.

* Central Limit Theorm:-

Image you have class of Students and you are interested in their heights. you measure each Student, and you notice that the height vary. Some Students are tall and Some are Short and distribution is not perfectly uniform.

Random Samples - Instead of measuring the entire class at once, you decide to take Small random groups of Students and Calculate the average height for each group.

Repeat the Process:-

You do this over and again, Creating many different Sample groups and you calculate the average height for each group.

Date: __/__/__

and may also refer you others. This process continue and your sample size grows.

4. Continue until you reach your goal:

You keep going until you have enough participant or until you have reached a point where you are satisfied with the diversity or information gathe

~~Explain Estimation~~.

Cluster - 3/5/22 (2020)

Cluster-sampling - Imagine you have a box of Colorful candies. you decide to grab a handful of smaller group of candies.

1. Divide into Group (cluster):- You look at your box of candies and notice they are already grouped into different bags. Each bag like a cluster of candies.

2. Pick a fewer ~~both~~ bags :- Even you check ~~bags~~ every single candy. You randomly grab a few bags from the box.

3. Check the bags you pick:- Now you open each bag and ~~candies~~ check candies inside. you are not looking at every candy indiviseu you just focus only the candies in the bag you picked.

#learnthesmarterway

Standard Error → σ is sample of Standard deviation

$SE = \frac{\sigma}{\sqrt{n}}$ n is number of Sample.

Date: __/__/__

Conditions of Chi-Square Test :-

(1) The two Sets data must be based on the Same Sample Size.

(2) Each Cell in the data contains the observed or expected count of five or large ?

MCQ :- The Test is used as test of Goodness of fit.

A

Snowball Sampling :-

Suppose you want to make a Snowball and ~~Snow~~ then roll it around.
As it roll, it pick up more Snow and get bigger.

1. Start with Seed :- You begin with Small group of people or an initial participant, who meets your criteria.

2. Roll the Snowball :- You ask this initial participant to find more participant.
They refer you might fit your Study.
Now your Sampling is growing, just like the Snowball rolling.

3. Keep Rolling :- Each new Participant you find become part of your study. ~~[illegible]~~ ~~[illegible]~~.

$$\chi^2 = \Sigma \frac{(O-E)}{E}$$

Sampling
Types
Central Limit Theorem

H_0 : birth of child is not associate with any day of the week or it is independence of the day of any week.

H_1: birth of child is associate with days of the week.

$$\chi^2_{(k-1)} = \sum_{i=1}^{k} \frac{(O_i - e_i)^2}{e_i}$$

e_i is expected value
O is the observe value

e_i is expected value
Or we can say O_i is the observed frequency
e_i is the expected frequency for each of the seven days of the week and n is the total number of observation.

$e_i = n \times p_i$
probability of happening of any event.

Observed and Accepted frequencies.

O_i	e_i	$O_i - e_i$	$(O_i - e_i)^2$	$(O_i - e_i)^2 / e_i$
184	150	34	1156	7.71
148	150	-2	4	0.03
145	150	-5	25	0.17
153	150	3	9	0.06
150	150	0	0	0
154	150	4	16	0.11
116	150	-34	1156	7.71
			$\chi^2 =$	15.77

145-150
-5

$\chi^2 = 15.77$.

-35-35
-10

Date: __/__/__

Karl Pearson coefficient of Correlation.

Correlation is a Statistical tool that help to measure and analyse the degree of relationship between two variable.

---

* Chi Square Goodness of Fit Test for Mean and Test of independence.

* chi-square Test (χ^2 Test): use

a) As a test of Significance for association / dependence :-

- H_0 : There is no association or dependence of one factor on the other.
- H_1 : There is Significant association or dependence of one factor on the other.

b) As a test of Goodness of Fit :-

- How good the observed frequencies fit a Specified assumption, patter or distribution.

---

* From a hospital record, the following data was obtained about the births of new born babies on various days of the week during past year.

Monday	Tue	Wed	Thur	Fri	Sat	Sun	Total
184	148	145	153	150	154	116	1050

[Here Sunday falls but Monday its peak up]

Possion Ratio:

Date: __/__/____

$$\boxed{P(x) = \frac{\lambda^k e^{1-\lambda}}{k!}}$$

e is the Euler number $(e = 2.71828......)$
K is the number of occurance.
K! is the factorial of K.

So, the positive real number λ is equal to it the Expected value of x and also its variance.

$$\lambda = E(x) = Var(x)$$

* The Equation can be adapted if, instead of the advantage number of event (λ), we are given a time rate for the events (r) to happen.

$\lambda = rt$ (Showing r number of events per unit time)

$$P(k \text{ event in interval } t) = (rt)^r e^{1-\lambda} \div k! \div k!$$

If the number of times an event occurs is an interval and k can take value 0, 1, 2, ... the occurance of one event does not affect the probability that a second event will occur.