Module 3

Sampling Distributions

Population

In ~~If~~ any statistical distribution ~~that~~ we are interested in studying the various characteristics of individuals (items) of a particular group. This group of individuals under study is known as population. Any population can be considered as the set of ~~observations~~ admissible value of a random variable. The distribution of this random variable is called the distribution of population. ~~If we~~

Eg: If we want study the expenditure habit of the family in a city then the population will consist of whole the household in that city.

A population contains finite no. of objects or items is known as finite population.

Eg. ~~A~~ students in a college → population in a city

Population having infinite and discrete no. of objects or with the no. of objects so large, termed as ~~discrete~~ infinite population.

Eg:- The population of temperature at the various point of the atmosphere.

Sampling

A finite subset of population, selected from it due to the objective of investigating its characteristics is called a sample. A sample is a representative part of population.

Eg:- When we want to study the life of electrical bulbs produce by a company is select some of bulbs and study their depth of life.

Large and small samples

When the sample space is more than 30 the sample is known as large sample otherwise small sample.

Statistic and Parameters

Any measure (function) calculated on the basis of population values is called parameter.

Eg:- population mean (μ), population variance (σ^2), population standard deviation (σ), population correlation coefficient (ρ - rau)

Any measure (function) calculated on the basis of sample values is called statistics.

Eg:- Sample mean ($\bar{x}$), sample variance (s^2), sample SD (s), sample correlation coefficient (r).

Sampling Distribution

Let $x_1, x_2, x_3 \dots x_n$ be a random sample taken from the population under investigation. We can consider the random observations as independent random variables ~~follow~~ following the same distribution of the population.

Let, $t = g(x_1, x_2, \dots x_n)$ be a function of this random variable, ~~Rv~~ also a random variable

The probability distribution 't' is called sam sample distribution of t.

In other words Listens why by a sample variable we mean the distribution of a statistics. if

If 't' is a statistic statistic is sample distribut belonged as $f(g)$

1. Find the pobility that there number of heads lee in the range 185 and 220 when fair coin is tossed 400 times.

0.9104

Hint : n is large $P = q$

$$\boxed{\begin{array}{l} x \sim N(m, \sigma^2) \\ \bar{x} \sim N(m, \frac{\sigma^2}{\sqrt{n}}) \end{array}}$$

$E(\bar{x}) = m$

$V(\bar{x}) = \sigma^2/\sqrt{n}$

Standard Error (SE)

If 't' is a statistic with sampling distribution $f(t)$, then standard error (SE) of P is given by

$$\boxed{SE(t) = \sqrt{v(t)}}$$

where, $v(t) = Et^2 - (E(t))^2$

standard Error of sample mean

suppose, the population variance is σ^2

$x_1, x_2 \dots x_n$ are sample points

$$\bar{x} = \frac{x_1 + x_2 + \dots x_n}{n}$$

$$V(\bar{x}) = V\left(\frac{x_1 + x_2 + \dots x_n}{n}\right)$$

$$= V\left(\frac{x_1}{n} + \frac{x_2}{n} + \dots \frac{x_n}{n}\right)$$

$$= V\left(\frac{x_1}{n}\right) + V\left(\frac{x_2}{n}\right) + \dots + V\left(\frac{x_n}{n}\right)$$

$$= \left(\frac{1}{n}\right)^2 V(x_1) + \left(\frac{1}{n}\right)^2 V(x_2) + \dots \left(\frac{1}{n}\right)^2 V(x_n)$$

$$= \frac{\sigma^2}{n^2} + \frac{\sigma^2}{n^2} + \dots \frac{\sigma^2}{n^2}$$

$$= n\frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$SE(\bar{x}) = \sqrt{V(\bar{x})}$$

$$= \sqrt{\frac{\sigma^2}{n}}$$

$$= \frac{\sigma}{\sqrt{n}}$$

$$\boxed{SE(\bar{x}) = \frac{\sigma}{\sqrt{n}}}$$, where σ is the population SD.

✓ **Sampling Distribution of sampling mean**

$$E(\bar{x}) = E\left(\frac{x_1 + x_2 + \cdots x_n}{n}\right)$$

$$= E\left(\frac{x_1}{n} + \frac{x_2}{n} + \cdots \frac{x_n}{n}\right)$$

$$= E\left(\frac{x_1}{n}\right) + E\left(\frac{x_2}{n}\right) + \cdots E\left(\frac{x_n}{n}\right)$$

$$= \frac{1}{n}E(x_1) + \frac{1}{n}E(x_2) + \cdots \frac{1}{n}E(x_n)$$

$$= \frac{m}{n} + \frac{m}{n} + \cdots \frac{m}{n}$$

$$= n \cdot \frac{m}{n}$$

$$= m$$

$$\boxed{\begin{aligned} E(\bar{x}) &= m \\ V(\bar{x}) &= \frac{\sigma^2}{n} \end{aligned}}$$

Consider a normal ~~distribution~~ population with mean (m) and variance (σ^2). Then the sample mean $\bar{x}$ has the normal distribution with mean (m) and variance (σ^2) ie, Let $x_1, x_2, \ldots x_n$ be a random sampling from,

$$N(m, \sigma^2)$$

Then,

$$\bar{x} = \frac{x_1 + x_2 + \cdots x_n}{n} \sim N\left(m, \frac{\sigma^2}{n}\right)$$

$$z = \frac{\bar{x} - m}{\left(\frac{\sigma}{\sqrt{n}}\right)} \sim N(0,1) \text{ or}$$

$$= (\bar{x} - m) \times \frac{\sqrt{n}}{\sigma}$$

$$= \frac{\sqrt{n}(\bar{x} - m)}{\sigma}$$

→ Distribution of ~~function~~ statistics function of samples or sample points
– Sample Distribution

eg: $\bar{x}$

→ Normal + Normal = Normal 49

~~x+y~~ (

→ $x+y \sim N(m_1+m_2,\ \sigma_1^2+\sigma_2^2)$

→ $x-y \sim N(m_1-m_2,\ \sigma_1^2+\sigma_2^2)$

Examples

1) $t = \frac{\bar{x}-m}{\sigma} \sim N(0,1)$

2) $t = \frac{\bar{x}_1-\bar{x}_2-(m_1-m_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$

where, $\bar{x}_1$ and $\bar{x}_2$ be the sample means of size n_1 & n_2 from the population $N(m_1, \sigma_1^2)$ and $N(m_2, \sigma_2^2)$ respectively.

Uses of Sampling distribution of 'z'

χ^2 - Distribution (Chi-Square Distribution)

It is another sample distribution. A continuous random variable chi-square (χ^2) is said to follow chi-square (χ^2) distribution if,

$$f(\chi^2) = \frac{\left(\frac{1}{2}\right)^{n/2} e^{-\chi^2/2} (\chi^2)^{n/2-1}}{\Gamma\frac{n}{2}} \quad 0<\chi^2<\infty$$

$\boxed{\Gamma = \text{Gamma}}$

we write, $\chi^2 \sim \chi^2(n)$

If 'z' follows a standard normal distribution, then z^2 follow chi-square (χ^2) distribution with one degree of freedom ($n=1$)

Since,

1 ~~Since~~ $\left(\frac{\bar{x}-m}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(1)$

If $z \sim N(0,1)$

then, $z^2 \sim \chi^2(1)$

(Contd.) 2) $\frac{ns^2}{\sigma^2} = \frac{ns^2}{\sigma^2} \sim \chi^2(n-1)$ $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$

2) **Additive property:**

If chi square 1 (χ_1^2) & χ_2^2 have independent χ^2 distributions with n_1 and n_2 degrees of freedom respectively. Then χ_1^2 & χ_2^2 will have a χ^2 distribution with $n_1 + n_2$ degrees of freedom. That is,

If, $\chi_1^2 \sim \chi^2(n_1)$ and
$\chi_2^2 \sim \chi^2(n_2)$

and they are independent, then $\chi_1^2 + \chi_2^2$ follows,

$$\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$$

Student's t- distribution

A continuous random variable assuming values from $-\infty$ to $+\infty$ with probability distribution function equal to,

$$f(t) = \frac{1}{\sqrt{n}\,\beta(\frac{1}{2}, \frac{n}{2})}\left(1 + \frac{t^2}{n}\right)^{\left(\frac{n+1}{2}\right)}, \quad -\infty < t < \infty$$

't' statistics

If the random variables 'z' follows, $N \sim (0,1)$ and Y follows $\chi^2(n)$ and if z and Y are independent then the statistics defined by

$$\boxed{t = \frac{z}{\sqrt{\frac{Y}{n}}}}$$ follows students 't' distribution with 'n' degrees of freedom.

Let X be a ~~capital~~ X $N(\mu, \sigma^2)$

$\bar{x}$ and s^2 be the sample mean and sample variance of a random samples of size.

47

Then, $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ (z = standard normal)

$$Y = \frac{ns^2}{\sigma^2} \sim \chi^2(n-1) \quad (Y = \chi^2\text{-distri})$$

$$t = \frac{Z}{\sqrt{\frac{Y}{n}}} = \frac{\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{\left(\frac{ns^2}{\sigma^2}\right)}{n-1}}}$$

$$= \frac{\cancel{\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}}}{}$$

$$= \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \times \frac{n-1}{\cancel{\left(\frac{ns^2}{\sigma^2}\right)}}$$

$$= \frac{\bar{x}-\mu}{\sqrt{\frac{\left(\frac{ns^2}{\sigma^2}\right)}{n-1}}} = \frac{\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}}{\frac{\left(\frac{\sqrt{n}s}{\sigma}\right)}{\sqrt{n-1}}} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \times \frac{\sqrt{n-1}}{\frac{\sqrt{n}s}{\sigma}}$$

$$= \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \left(\sqrt{n-1} \times \frac{\sigma}{\sqrt{n}s}\right)$$

$$= \frac{\bar{x}-\mu}{} \left(\sqrt{n-1} \times s\right)$$

$$= \frac{\bar{x}-\mu}{s/\sqrt{n-1}}$$

$$\boxed{\frac{\bar{x}-\mu}{s/\sqrt{n-1}}} \text{ Ans.}$$

Uses of 't' Distribution

The variable 't' is a ~~static~~ statistic and it is used in many tests of hypothesis. Those tests are known as 't' test. And use,

a) to test the given population mean when sample is a small

b) To test whether the 2 samples have same mean when samples are small.

c) To test whether there is difference in the observations of the 2 dependent samples.

d) To test the significance of population correlation coefficient.

Uses of χ^2 Distribution

1) To test the given population variance when sample is small.

2) To test the goodness of fit between ~~expected~~ observed and expected frequencies.

3) To test the independence of 2 attributes.

4) To test the homogeneity of data.

'F' - Distribution

'F' stands for Fisher - Father of modern statistics

A random variable 'F' is said to follow 'F' distribution if it probability function is,

$$f(F) = \begin{cases} \dfrac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} F^{\frac{n}{2}-1} \left(1+\frac{n_1}{n_2}F\right)^{-\left(\frac{n_1+n_2}{2}\right)}}{\beta\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} & 0 < F < \infty \\ 0, & \text{otherwise,} \end{cases}$$

We denote, $F \sim F(n_1, n_2)$

Uses of 'F' - Distribution

F-statistics is used for test of hypothesis. The test conducted on the basis of 'F' statistic is called 'F' test. F-test can be used to

1 - Test the equality of variance of 2 populations when samples are small.

2 - Test the equality of means of 3 or more populations. ANOVA (Analysis of Variance)

'F' - Statistics

Let 'U' follows $\chi^2(n_1)$ and 'V' follows $\chi^2(n_2)$ and U & V are independent.

Then, $F = \dfrac{U/n_1}{V/n_2} \sim F(n_1, n_2)$

Sample 1 | sample 2

$$U = \frac{n_1 s^2}{\sigma_1^2} \sim \chi^2(n_1-1) \quad \Big| \quad V = \frac{n_2 s^2}{\sigma_2^2} \sim \chi^2(n_2-1)$$

$$F = \frac{\left(\frac{n_1 s^2}{\sigma^2}\right) / n_1-1}{\frac{n_2 s^2}{\sigma_2^2} / n_2-1} \sim F(n_1-1, n_2-1)$$

Relations b/w Normal, chi-square, t & F Distr

→ when x follows a normal distribution with mean $= \mu$ & SD $= \sigma$. Then,

$z = \dfrac{x-\mu}{\sigma}$ follows standard normal.

→ when $z_1, z_2, \dots z_k$ are k standard normal variables then,

$\sum_{i=1}^{k} Z_i^2$ follows chi-square distribution

with k ~~decreasable~~ of freedom

→ If Z follows standard normal distribution and Y follows chi-square with k decrease of freedom then,

$\frac{Z}{\sqrt{\frac{Y}{k}}}$ ~ t follows 't' distribution with k decrease of freedom.

→ If Y_1 & Y_2 are a chi-square variance with n_1 & n_2 decrease of freedom. Then,

$\frac{Y_1/n_1}{Y_2/n_2}$ ~ follows 'F' distribution with n_1 & n_2 decrease of freedom.

Questions

1. For a continous random variable $P(a < x \leq b)$ is,

A) $F(b) - F(a)$
B) $F(a) - F(b)$
c) $F(b+h) - F(a-h)$
d) $F(b+h) - F(a+h)$

2- Sampling variance of mean based on a sample of size n is,

A) ~~$\sigma/\sqrt{n}$~~ $\sigma/\sqrt{n}$ B)
E) σ^2/n D)

3- For a poisson distribution with parameter 4, variance is,

a) 2 c) 16
b) 4 d) 8

4- For a binomial distribution ~~the~~ mean 6, & variance is 3 then n, is - ___

5. If x follows standard normal then x^2 follows

A) F-distribution
B) Chi-square distribution
C) Normal Distribution
D) t - Distribution

6- Binomial distribution with parameter p is symmetric when,

A - $P = 1/2$

Scanned by CamScanner

A) $P < 1/2$ c) $P = 1/2$

b) $P > 1/2$ D) $P \geq 1/2$

7. 5^{th} central moment of normal distribution is

A) 1 c) $5\sigma^2$

b) 0 D) $\mu_5 + 3\sigma^2$

Answers

6 $F(b) - F(a)$

2. σ^2/n

3. 4

4. $np = 6$

$npq = 3$

$6q = 3$

$q = 3/6 = 1/2$

$p = 1 - 1/2 = 1/2$

$n \times 1/2 = 6$

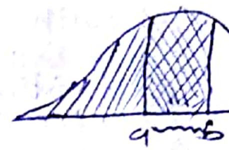
$n = 6/\tfrac{1}{2} = 6 \times 2 = 12$

5. chi-square distribution

6. $P = 1/2$

7. odd order central moment is zero

b) = 0



1. $F(x) = P(X \leq x)$

$F(b) = P(X \leq b)$

$F(a) = P(X \leq a)$

$= F(b) - F(a)$

Imp

Central Limit Theorem

Let $x_1, x_2 \dots x_n$ be n independent random variables ~~Samples~~. Let all have same distribution, independent identical same mean. say μ and same standard deviation, σ, Then the mean of all these variables, $\frac{x_1 + \dots + x_n}{n}$ follows a normal distribution with mean μ and $SD = \frac{\sigma}{\sqrt{n}}$ when n is large.