**Simple random sampling from finite population**

Sampling from a finite population means selection of some number of items (say, ) from an aggregate of items (say, ) and .

A simple random sample of size consists of individuals chosen from the population in such a way that every possible set of individuals has an equal chance to be the sample actually selected.

There can be two distinct types of simple random sampling in case of finite population. These are –

* Simple random sampling with replacement (SRSWR) and
* Simple random sampling without replacement (SRSWOR)

**Simple random sampling with replacement (SRSWR):** In this case, the units of the sample are drawn from the population one by one, after each drawing the individual selected being returned to the population in such a way that at each drawing each of the members of the population gets the same probability of being selected.

* Probability of selection of a sample of size of a particular sequence
* Probability of selection of a sample of size (where sequence of items are ignored)

**Simple random sampling without replacement (SRSWOR):** If the members of the sample are drawn one by one but the member obtained at any drawing is not returned to the population and if at each stage every remaining unit of the population (at drawing each of the remaining units) is given the same probability of being included in the sample, then we have simple random sampling without replacement.

* Probability of selection of a sample of size of a particular sequence

* Probability of selection of a sample of size (where sequence of items are ignored)

The above probability can be estimated using a different approach as discussed below:

* Obviously, from a population of items a sample of size can be taken in different ways.
* Therefore, the probability of collecting a sample of size is .

Hence,***if one takes individuals all at a time from the population, he/she will still have simple random sampling without replacement.***

* Probability of any specified individual being selected at any specified drawing, Prob ( individual is selected at drawing)

**Some important terminologies: statistic, parameter, sampling fluctuation, sampling distribution and standard error**

Suppose, a population consists of members and one character of importance is a variable . Let is the value of the variable for the member of the population.

Note that a ***random sample of size***  is essentially collection of  random variables , , ..., that are mutually independent.

The observed values () can be regarded as the value assumed by the random variables (). The set of values , ,..., is called a realization of the sample.

**Statistic:** A measure of population distribution of (e.g. mean, standard deviation etc.), calculated on the basis of sample values, i.e. , ,..., , is called a statistic.

In general, statistic is a function of the sample observations, e.g. sample mean () and standard deviation () are statistic, where

and

**Parameter:** A measure of population distribution of (e.g. mean, standard deviation etc.), calculated on the basis of population values of , is called a parameter of the distribution. For example, and are parameters, where

and

**Sampling fluctuation:** Since the sets of population members included in different samples from the same population may be different, the values of the statistic itself are liable to vary from one sample to another. These differences in the values of a statistic are called sampling fluctuation.

**Sampling distribution:** If a number of samples, each of size , are taken from the same population and if for each sample the value of the statistic is calculated, a series of values of the statistic will be obtained. If the number of samples is large, these may be arranged into a frequency table. The frequency distribution of the statistic that would be obtained if the number of samples each of size were infinite, is called the sampling distribution of the statistic.

**Standard error:** Like any other distribution, a sampling distribution may have its mean, standard deviation and moments of higher orders. The standard deviation of a sampling distribution is designated as the standard error of the statistic.

**Sampling distribution of sample mean: Measures of central tendency and dispersion**

**Case 1:Random sampling with replacements (SRSWR)**

Suppose a random sample of size is drawn from a population of size .

.....................................................(1)

Now,

for each

(Since & are independent)

Thus, from equation (1) we get,

Standard error of ,

Therefore, in case of SRSWR, we find that

* Standard error of ,

**Case 2: Random sampling without replacement**

Here too can take one of the values , with the same probability . Therefore, as before here too, for each,

However, the covariance terms need special attention. Here for ,

Hence,

Thus,

, since # terms =

Standard error of ,

Therefore, in case of SRSWOR, we find that

* Standard error of ,

Note that

* In both the cases (SRSWR & SRSWOR), the standard error decreases with increasing .
* The standard error of the mean in sampling without replacement is, however, smaller than that in sampling with replacements.
* But the difference becomes negligible if is very large compared to .
* In SRSWOR, the standard error of the sample mean vanishes if , which is expected because the sample mean now becomes a constant, i.e. the same as the population mean. However, this is not the case with sampling with replacement.

**Sampling distribution of sample proportion: Measures of central tendency and dispersion**

Suppose in a population of members, there are  ***members with a character A*** and  ***members with the character not-A***. Then is the proportion of members in the population having the character A.

Let a random sample of size be drawn from the population, and let be the numbers of members in the sample having the character A. Then the sample proportion is . To find the expectation and standard error of the sample proportion, we adopt the following procedure.

We assign to the member of the population the value , where

In this way, we get a variable , which has

* Population mean,

, and

* Population variance,

, since

, where

Similarly, we assign to the member of the sample the value , where

Then, the sample mean of the variable is

**Case 1:Random sampling with replacements (SRSWR)**

We know that in case of random sampling with replacement (SRSWR),

(1)

(2)

Now replacing by , by and by in these two equations, we get the expectation of sample proportion and standard error of sample proportion as follows:

(3)

(4)

**Case 2:Random sampling without replacements (SRSWOR)**

Similarly, we know that in case of random sampling without replacement (SRSWOR),

(5)

(6)

Replacing by , by and by in these two equations, we get the expectation of sample proportion and standard error of sample proportion as follows:

(7)

(8)

**Random Sampling from an Infinite Population**

* There will be many cases where the population has to be considered infinite and hypothetical. For examples:
* Suppose one is interested in knowing ***whether a given coin is unbiased***. The population then comprises the whole set of throws that may be made with the coin. This is clearly infinite and hypothetical.
* Suppose one is interested in the ***average life (in hours) of a type of blub*** ***produced by a bulb factory***. Here the population to which the average relates includes not only the blubs that the factory has in stock, but also all blubs that it produced in past and sold out and all those that it may produce in future. Here too the population is infinite and hypothetical.
* The whole idea of an infinite population is clearly quite abstract. One way to think of it is to consider sampling from a finite population, and increasing the size of the population: suppose that the population size *N* tends to infinity.
* Sampling from an infinite population is handled by regarding the population as represented by a distribution. ***A random sample from an infinite population is, therefore, considered as a random sample from a distribution***.
* This means that some values more likely in the random sample than others, according to the shape of the distribution. The underlying distribution can be thought of as the distribution of some ***random variable X***.
* A random variable *X* can be considered as a numerical description of the outcome of a statistical experiment or random phenomenon. The [***probability distribution***](https://www.britannica.com/science/distribution-function)  for a random variable describes ***how the probabilities are distributed over the values of the random variable***.
* The random variable that may assume ***only a finite number or an infinite sequence of values*** is said to be **discrete**. For instance, a random variable representing the number of automobiles sold at a particular dealership on one day would be discrete.
* The random variable that may assume ***any value in some interval on the real number line*** is said to be **continuous**. For instance, a random variable representing the weight of a person in kilograms would be continuous.
* For a discrete random variable , the probability distribution is defined by a probability mass function, denoted by . This function provides the probability for each value of the random variable. For a continuous random variable , the probability distribution of is defined by a probability density function, also denoted by .

**Meaning of random sample of size from an infinite population**

* A ***random sample of size***  from an infinite population or from a probability distribution of is defined to be  random variables , , ..., that are mutually independent and have the same distribution as .
* The observed values () can be regarded as the value assumed by the random variables (), where , , ..., are independent random variables with common pmf/pdf as of . The set of values , ,..., is called a realization of the sample. The possible values of the random vector (, , ..., )can be regarded as points in the ***sample space***.

**A simple random sample from a very large finite population is approximately the same as a random sample from an infinite population**

* If we draw two numbers at random, without replacement, ***from a population consisting of the integers 1, 2, 3, 4 and 5***, the second number is clearly not independent of the first number. For example, if we define A= first number is 3 and B=second number is 4, then P(A)=1/5, but P(B/A)=1/4, and since these probabilities are not equal, the two events are not independent.
* On the other hand, if we draw two numbers at random, without replacement, ***from a population consisting of the integers 1, 2, 3, …, 10000***, then the corresponding events have the following probabilities: P(A) = 1/10000, but P(B|A) = 1/9999. These are different, so the two events are not independent, but they are very close, so the events in this case are approximately independent: Pr(A) ≈ Pr(B|A).
* Now, ***imagine a huge population of marbles, say*** . Now the population is so vast that each time a marble is selected, it will effectively be a selection from this distribution, no matter what other marbles have been selected. That is, each observation has the same distribution independently of the other observations.

**Joint probability mass/density function of a random sample**

If is the of , then the joint probability mass/density function of a random sample is given by

Using the above knowledge, we can derive some common sampling distribution.

**Sampling distribution of sample total: Poisson parent**

Suppose and are distributed independently in the Poisson form with parameters and respectively. The sum can then take the value 0,1, 2,...

Now,

i.e. () itself is a Poisson variable with parameter ().

It immediately follows that

* If , ,..., are independently distributed Poisson variables with parameters ,..., then the sum is also a Poisson variable with parameter .
* If , ,..., are a random sample from a Poisson distribution with parameters , then sampling distribution of the statistic is also a Poisson with parameter .

**Sampling distribution of sample total: Binomial parent**

Suppose and are distributed independently in the binomial form with parameters and respectively. The sum can then take the value 0,1, 2,...., .

Now,

The sum is nothing but the sum of products of the coefficients of in and coefficients of in for varying , and equals the coefficient of in , which is equals to .

Thus, , which implies that () follows binomial distribution with parameters () and .

It immediately follows that

* If , ,..., are independently distributed binomial variables with parameters ; ;..., , then the sum is also a binomial variable with parameters) and .
* If , ,..., are a random sample from a binomial distribution with parameters and , then sampling distribution of the statistic is also binomial with parameter and .

**If is normally distributed with mean and variance , then , where , is also normally distributed with mean and variance .**

***Proof:*** Let us denote the pdf of and by and respectively.

If , we have,

If , we similarly have,

Therefore, combining the two results, we get,

This is nothing but the density function of a normal variable with mean and variance . Therefore, the above theorem is proved.

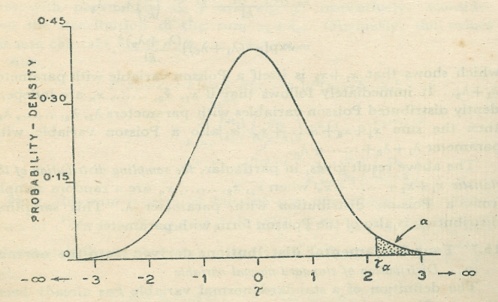
**Four fundamental distributions derived from the normal distribution**

# *Standard normal distribution*

Let , where is a normal variable with mean and standard deviation . Then itself is a normal variable which has mean zero and standard deviation unity. Such a normal variable is called a standard normal variable. The probability density function of the standard normal distribution is

,

* Mean,
* Variance,
* The upper point of standard normal distribution is denoted as , i.e. P[Z>]=, and the lower point of standard normal distribution is denoted as , i.e. P[Z<]=.
* Because of the symmetry of the distribution about zero, = .
* If is a standard normal variable, then is a normal variable with mean and variance .

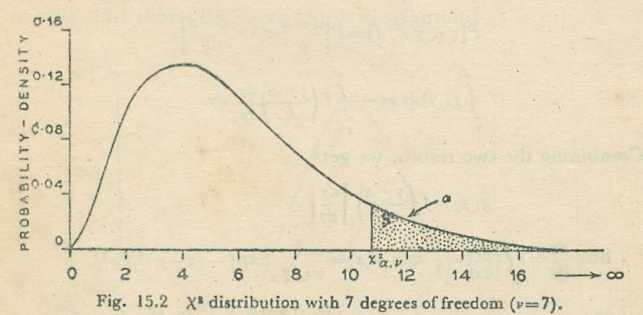


* ***distribution***

Let , , , ..., be mutually independent standard normal variables. Then the sum of their squares, i.e.  is called a **(**chi-square) variable with degrees of freedom (). It has the probability density function

,

* Mean, i.e.
* Variance =
* The distribution is always positively skew, and for there is a unique maximum at .
* If and be two independent variables distributed as with df equal to and respectively, then + is distributed as with df equal to + .
* For large , is approximately normally distributed with mean and standard deviation 1.
* The upper -point of distribution with is denoted as , i.e. , and the lower -point of distribution with is denoted as , i.e. ].

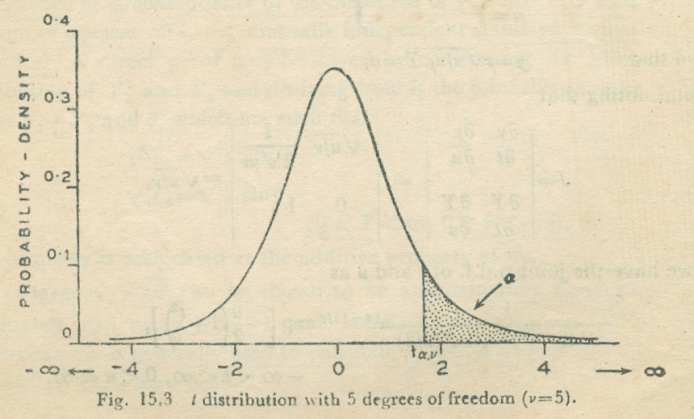


* ***distribution***

# If be a standard normal variable and a chi-square variable with , distributed independently of , then the new variable is called a variable with degrees of freedom. It has the probability density function

,

* Mean, i.e.
* Variance, for
* Like the standard normal distribution, the t distribution is symmetrical about . But it has , i.e. more peaked than the standard normal distribution with the same standard deviation.
* The upper -point of distribution with is denoted as , i.e. , and the lower -point of distribution with is denoted as , i.e. ].
* Because of the symmetry of the distribution about zero, = .



* ***distribution***

# Let and be independently distributed as s with and degrees of freedom, respectively. The random variable is then called an variable with (,) degrees of freedom. The probability density function of distribution is

,

* Mean, i.e. for
* Variance, , for
* It can be shown that
* The F distribution is highly positively skew.
* The upper -point of distribution with is denoted as , i.e. , and the lower -point of distribution with is denoted as , i.e. ].
* and

