Statistics : The basic problem of statistics is to determine the probability of an event A connected with a random experiment E. Although there are several definitions of probability of an event, viz., (I) Classical definition of probability, (II) Axiomatic definition of probability and (III) Frequency interpretation of probability but no one is helpful to determine the probability of an event A connected with E in any real life problem. So, in statistics, we want to estimate $P(A)$, probability of A connected with E. Suppose p_d is an estimate of $P(A)$. Then our task is to make or consider a hypothesis on $P(A)$ with respect to its estimated value p_d.

The basic problem of statistics may also be described as follows:

Let X be a random variable defined on an event space S connected with a random experiment E. The functional form of the probability distribution function of X may be known or unknown. If the functional form of the probability distribution function of X is $F_X(x)$, then the form of $F_X(x)$ may be known but it contains some unknown parameters. Then the problem of statistics is to estimate the values of the unknown parameters and finally we shall consider the hypothesis on the estimated values of the unknown parameters with respect to their exact values. On the other hand, if the functional form of $F_X(x)$ is not known, then our task is to make a hypothesis regarding the functional form of $F_X(x)$ and we want to determine the parameters associated with the approximated functional form of $F_X(x)$.

Some Important Theorems of Probability

Theorem : If X is a $\Gamma(\frac{n}{2})$ variate, then $Y = 2X$ is a χ^2- variate with n degrees of freedom (df) and conversely if Y is a χ^2- variate with n df then $X = \frac{Y}{2}$ is a $\Gamma(\frac{n}{2})$ variate.

Theorem-II : If $X_1, X_2, \dots, X_n$ are n mutually independent standard Normal variates then $X_1^2 + X_2^2 + \dots + X_n^2$ is a χ^2- variate with n degrees of freedom.

✓ **Theorem – III**: Let $X_1, X_2, \ldots, X_n$ be n mutually independent standard normal variates. If

$$a_{11}X_1 + a_{12}X_2 + \cdots + a_{1n}X_n = \sum_{j=1}^{n} a_{1j}X_j,$$

$$a_{21}X_1 + a_{22}X_2 + \cdots + a_{2n}X_n = \sum_{j=1}^{n} a_{2j}X_j,$$

$$\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots$$

$$a_{m1}X_1 + a_{m2}X_2 + \cdots + a_{mn}X_n = \sum_{j=1}^{n} a_{mj}X_j,$$

or equivalently as

$$a_{\beta 1}X_1 + a_{\beta 2}X_2 + \cdots + a_{\beta n}X_n = \sum_{j=1}^{n} a_{\beta j}X_j$$

for $\beta = 1(1)m$ are $m\ (<n)$ linear combinations such that their coefficients (a_{ij}) satisfy orthogonality conditions:

$$\sum_{\alpha=1}^{n} a_{i\alpha} a_{j\alpha} = \delta_{ij} \quad \forall\ i, j = 1, 2, \ldots, m$$

then $Q = \sum_{i=1}^{n} X_i^2 - \sum_{\beta=1}^{m} (a_{\beta 1}X_1 + a_{\beta 2}X_2 + \cdots a_{\beta n}X_n)^2$

is χ^2 variate with $(n-m)$ d.f. and Q is independent of the given linear combinations.

Theorem IV: If

1. X is a standard normal variate,
2. χ^2 is a χ^2- variate with n df,
3. X and χ^2 are indpendent,

then

$$y = \frac{X}{\sqrt{\chi^2/n}} = \sqrt{n}\,\frac{X}{\sqrt{\chi^2}}$$

is a t- variate with n df.

Theorem V: $\boxed{\text{Tchebychett's Inequality}}$ → If

X is a random variable having finite variance σ^2 and mean m, then for any $\varepsilon > 0$,

$$P(|X - m| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

or

$$P(|X - E(X)| > \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

or

$$1 - P(|X - E(X)| \leq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

$$\Rightarrow P(|X - E(X)| \leq \varepsilon) \geq 1 - \frac{Var(X)}{\varepsilon^2}$$

Convergence in Probability: A sequence of random variables $X_1, X_2, \ldots, X_n, \ldots$ is said to converge in probability to a constant 'a' or a random variable 'X' if

$$\lim_{n\to\infty} P(|X_n - a| < \varepsilon) = 1 \quad \text{or} \quad \lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$$

or equivalently as

$$\lim_{n\to\infty} P(|X_n - a| \geq \varepsilon) = 0 \quad \text{or} \quad \lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$$

for any $\varepsilon > 0$.

In this case, we write

$$X_n \xrightarrow[mp]{} a \quad \text{or} \quad X_n \xrightarrow[mp]{} X \quad \text{as } n \to \infty$$

Theorem VI: Let $X_n \xrightarrow[mp]{} X$ and $Y_n \xrightarrow[mp]{} Y$ as $n \to \infty$ then

1. $X_n \pm Y_n \xrightarrow[mp]{} X \pm Y$

2. $cX_n \xrightarrow[mp]{} cX$, for any constant c

3. $X_n Y_n \xrightarrow[mp]{} XY$

4. $\dfrac{X_n}{Y_n} \xrightarrow[mp]{} \dfrac{X}{Y}$ provided $Y \neq 0$

Theorem VII : $\boxed{\text{Tchebycheff's Theorem}}$ → Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of random variables such that mean m_n and standard deviation σ_n of X_n exist for all n. If $\sigma_n \to 0$ as $n \to \infty$ then $X_n - m_n \xrightarrow[mp]{} 0$ as $n \to \infty$

Theorem VIII : Law of Large numbers → If $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of random variables such that

① $S_n = X_1 + X_2 + \cdots + X_n$ has finite mean M_n $\forall n$

② S_n has finite variance σ_n $\forall n$

③ $\sigma_n = O(n)$, i.e., $\frac{\sigma_n}{n} \to 0$ as $n \to \infty$

then

$$\frac{S_n - M_n}{n} \xrightarrow[mp]{} 0 \quad \text{as } n \to \infty$$

Population, Sample, Fake random variable, Different characteristics of a sample

Population: Let X be a random variable defined on an event space S connected with a random experiment E. So, any particular performance of E results a value of X and hence one can get a sequence of values of X for a sequence of repetitions of the random experiment E under uniform and identical conditions. So, a sequence of repetitions of E under uniform and idential conditions will give a sequence of values of X, the totality of which will be called the population of the random variable X connected with E.

Sample or Random Sample:

Let X be a random variable defined on an event space S connected with a random experiment E. Now a sequence of n repetitions of E under uniform and identical conditions will give a sequence of n observed values of X: $x_1, x_2, \ldots, x_n$. This '<u>n observed values of X</u>': $x_1, x_2, \ldots, x_n$ is known as a <u>sample of size 'n'</u> drawn from the population of the random variable X. Now if we perform another n repetitions of E under uniform and identical conditions, then, in general, we will get <u>another</u> n observed values of X: $x_1', x_2', \ldots, x_n'$. Thus if different samples of size n are drawn repeatedly under uniform and identical conditions the set of observed values of X will be different and in this sense, a sample is also known as random sample.

Fake Random Variable:

Let X be a random variable defined on an event space S connected with a random experiment E. Let us consider a sample : $x_1, x_2, \ldots, x_n$ of size n drawn from the population of X. To describe the distribution (probability distribution) of the sample values : $x_1, x_2, \ldots, x_n$, we define a random variable $\mathring{x}$ which takes the sample values only each with probability $\frac{1}{n}$,

i.e.,

$$P(\mathring{x} = x_i) = \frac{1}{n} \quad \forall\ i = 1(1)n$$

The random variable $\mathring{x}$ is known as the Fake random variable corresponding to the given sample.

So, the probability distribution of the fake random variable is always discrete type irrespective of the nature of the distribution of the parent random variable X.

Emperical distribution:

The probability distribution function of the Fake random variable corresponding to a sample drawn from the population of a random variable is known as the emperical distribution or the distribution of the sample.

Theorem: The emperical distribution is the statistical image of the distribution of the population.

Proof: Let us consider a sample of size n $x_1, x_2, \ldots, x_n$ drawn from the population of the random variable X connected with a random experiment E. Therefore, the fake random variable $\mathring{x}$ corresponding to the above-mentioned sample values is given by

$\mathring{x} = \{x_1, x_2, \ldots, x_n\}$,

where

$P(\mathring{x} = x_i) = \frac{1}{n} \quad \forall\, i = 1(1)n.$

Let $F_X(x)$ be the probability distribution function of the parent random variable X whereas $G(x)$ is the probability distribution function of $\mathring{x}$

$\therefore\ G(x) = P(\mathring{x} \leq x) = \frac{r}{n}$,

where r is the number of sample values such that the observed value of each sample $\leq x$

$\therefore \frac{r}{n}$ is the frequency ratio of the event $X \leq x$ for sufficiently large n.

$\therefore F_X(x) = P(X \leq x) \approx \frac{r}{n}$ for sufficiently large n

$\Rightarrow F_X(x) \approx G(x)$ for sufficiently large n

$\Rightarrow$ Emperical distribution or distribution of a sample (or the distribution of the Fake random variable is the statistical image of the distribution of the population.

<u>Characteristics of sample:</u> The mean, variance moments, skewness, kurtosis etc. of the Fake random variable are known as different characteristics of sample. These characteristics are helpful to

understand the nature of the emperical distribution. Let us consider a sample of size n : $x_1, x_2, \ldots, x_n$ drawn from the population of the random variable X defined on an event space S connected with a random experiment E. Therefore, the Fake random variable $\mathring{X}$ is given by

$\mathring{X} = \{x_1, x_2, \ldots, x_n\}$,

where

$P(\mathring{X} = x_i) = \frac{1}{n} \quad \forall \, i = 1(1)n$.

<u>Sample mean</u> : Sample mean is denoted by $\bar{x}$ and defined by

$$\bar{x} = E(\mathring{X}) = \sum_{i=1}^{n} x_i \, P(\mathring{X} = x_i) = \sum_{i=1}^{n} x_i \times \frac{1}{n}$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

<u>Sample Variance</u> : This is generally denoted by s^2 and defined by

$$s^2 = Var(\mathring{x}) = E(\mathring{x} - E(\mathring{x}))^2$$

$$= E(\mathring{x} - E(\mathring{x}))^2 = E(\mathring{x} - \bar{x})^2$$

$$= \sum_{i=1}^{n} (x_i - \bar{x})^2 P(\mathring{x} = x_i)$$

$$= \sum_{i=1}^{n} (x_i - \bar{x})^2 \frac{1}{n}$$

$$= \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} (x_i^2 - 2\bar{x} x_i + \bar{x}^2)$$

$$= \frac{1}{n} \left[\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} 2\bar{x} x_i + \sum_{i=1}^{n} \bar{x}^2 \right]$$

$$= \frac{1}{n} \left[\sum_{i=1}^{n} x_i^2 - 2\bar{x} \sum_{i=1}^{n} x_i + n\bar{x}^2 \right]$$

$$= \frac{1}{n} \left[\sum_{i=1}^{n} x_i^2 - 2\bar{x}\, n\bar{x} + n\bar{x}^2 \right]$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n} \sum_{i=1}^{n} x_i \right)^2$$

<u>Kth order sample raw moment</u> : This is generally denoted by a_k and defined by

$$a_k = k\text{th order sample raw moment}$$
$$= k\text{th order raw moment of } \mathring{x}$$
$$= E(\mathring{x}^k)$$
$$= \sum_{i=1}^{n} x_i^k \, P(\mathring{x} = x_i)$$
$$= \sum_{i=1}^{n} x_i^k \left(\frac{1}{n}\right)$$
$$= \frac{1}{n} \sum_{i=1}^{n} x_i^k = \frac{x_1^k + x_2^k + \cdots + x_n^k}{n}$$

<u>Theorem</u> : $a_0 = 1$, $a_1 = \bar{x}$

Proof: $a_0 = \frac{1}{n} \sum_{i=1}^{n} x_i^0 = \frac{1}{n} \sum_{i=1}^{n} 1 = \frac{1}{n}(1+1+\cdots+1)$

$$= \frac{1}{n} \times n = 1.$$

$$a_1 = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$$

<u>k th order sample central moment</u>: This is generally denoted by m_k and defined by

m_k = k th order sample central moment

= k th order central moment of $\overset{\circ}{X}$

$$= E(\overset{\circ}{X} - E(\overset{\circ}{X}))^k = E(\overset{\circ}{X} - \bar{x})^k$$

$$= \sum_{i=1}^{n} (x_i - \bar{x})^k P(\overset{\circ}{X} = x_i)$$

$$= \sum_{i=1}^{n} (x_i - \bar{x})^k \left(\frac{1}{n}\right)$$

$$= \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^k$$

<u>Theorem</u>: $m_0 = 1$, $m_1 = 0$, $m_2 = s^2$

<u>Proof</u>: According to the definition,

$$m_k = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^k$$

$$\therefore m_0 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^0 = \frac{1}{n} \sum_{i=1}^{n} 1 = \frac{1}{n}(1+1+\cdots+1) = 1.$$

$$m_1 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^1 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})$$

$$m_1 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}) = \frac{1}{n}\left[\sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\bar{x}\right]$$

$$= \frac{1}{n}\left[n\bar{x} - (\bar{x} + \bar{x} + \cdots + \bar{x})\right]$$

$$= \frac{1}{n}\left[n\bar{x} - n\bar{x}\right] = 0$$

$$m_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = s^2$$

<u>Theorem</u>: $m_k = \sum_{i=0}^{k}(-1)^i \binom{k}{i} a_{k-i}\,\bar{x}^i$

<u>Proof</u>: According to the definition,

$$m_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^k = \frac{1}{n}\sum_{i=1}^{n}(-\bar{x} + x_i)^k$$

$$\Rightarrow m_k = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=0}^{k}\binom{k}{j} x_i^{k-j}(-\bar{x})^j$$

$$\Rightarrow m_k = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=0}^{k}\binom{k}{j} x_i^{k-j}(-1)^j\,\bar{x}^j$$

$$= \frac{1}{n}\sum_{j=0}^{k}\sum_{i=1}^{n}\underbrace{\binom{k}{j}(-1)^j\,\bar{x}^j}_{\text{depends on } j}\, x_i^{k-j}$$

$$= \frac{1}{n}\sum_{j=0}^{k}\binom{k}{j}(-1)^j\,\bar{x}^j\sum_{i=1}^{n} x_i^{k-j}$$

$$m_k = \frac{1}{n} \sum_{j=0}^{k} \binom{k}{j} (-1)^j \bar{x}^j \sum_{i=1}^{n} x_i^{k-j}$$

$$= \sum_{j=0}^{k} \binom{k}{j} (-1)^j \bar{x}^j \underbrace{\left(\frac{1}{n} \sum_{i=1}^{n} x_i^{k-j} \right)}_{\checkmark (k-j) \text{ th order raw moment}}$$

$$= \sum_{j=0}^{k} \binom{k}{j} (-1)^j \bar{x}^j a_{k-j}$$

$$= \sum_{j=0}^{k} (-1)^j \binom{k}{j} a_{k-j} \bar{x}^j$$

$$= \sum_{i=0}^{k} (-1)^i \binom{k}{i} a_{k-i} \bar{x}^i \quad [i \leftrightarrow j]$$

Skewness: Skewness is a measure of asymmetry of the emperical distribution of the Fake random variable.

Coefficient of skewness of $\mathring{x} = g_1 = \frac{m_3}{s^3}$

For normal distribution $m_3 = 0 \Rightarrow g_1 = 0$

For unimodal distributions, positive skew indicates that the tail is on the right whereas negative skew indicates that the tail is on the left side. A zero skew indicates that the tails on the both sides of the mean of the distribution. Again for positive skew, Mode < Median < Mean, for negative skew, Mean < Median < Mode whereas for zero skew, Mean = Median = Mode. There are several methods to describe skewness of a probability distribution but here we discuss the skewness with respect to the normal distribution.

Kurtosis: Kurtosis is a measure of tailedness of the probability distribution. It is measured by the coefficient of Kurtosis which is defined as

$$g_2 = \frac{m_4}{s^4}$$

For normal distribution, $g_2 = 3$.

Sampling

Let us consider a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of a random variable X connected with a random experiment E. This samble can be regarded as a random sample in the sense that if we draw another sample of size n from the population of X then this samble is not exactly same as the previous sample $x_1, x_2, \ldots, x_n$. Therefore, the first sample value can be regarded as an observed value of the random variable X_1, the second sample value can be regarded as an observed value of the random variable X_2, and so on, and finally, x_n can be regarded as an observed value of the random variable X_n where

X_1	X_2	...	X_n
x_1	x_2	...	x_n
y_1	y_2	...	y_n
z_1	z_2	...	z_n
t_1	t_2	...	t_n

① $X_1, X_2, \ldots, X_n$ are mutually independent because we have considered n independent repetition of E to get a sample of size n.

② The probability distribution function of each X_i is same as X for all $i = 1(1)n$ because the sample values are, actually, the observed values of the parent random variable X.

Therefore a sample $x = (x_1, x_2, \ldots, x_n)$ can be regarded as a observed value of the random variable $Y = (X_1, X_2, \ldots, X_n)$ where each X_i has the same distribution as X and $X_1, X_2, \ldots, X_n$ are mutually independent.

The probability distribution of Y

$= $ The joint probability distribution of $X_1, X_2, \dots, X_n$

$= F_Y(x_1, x_2, \dots, x_n)$

$= F_{X_1}(x_1)\, F_{X_2}(x_2) \cdots F_{X_n}(x_n)$ [As $X_1, \dots, X_n$ are mutually independen]

$= F_X(x_1)\, F_X(x_2) \cdots F_X(x_n)$ [As each X_i has the same distribution as $X \Rightarrow F_{X_i}(x) = F_X(x_i)$]

Sample point & Sample space :-

Let us consider a sample of size n: $x_1, x_2, \dots x_n$ drawn from the population of a random variable X connected with a random experiment E. Let $X_1, X_2, \dots, X_n$ are the random variables corresponding to the sample values $x_1, x_2, \dots, x_n$ respectively, where

① $X_1, X_2, \dots, X_n$ are mutually independent

② Each X_i has the same probability distribution as X.

Then the n dimensional random variable

$\vec{x} = (X_1, X_2, \dots X_n)$

is known as a sample point and the space generated by the sample points is known as the sample space. This sample space is also denoted by R^n. But it is not equal to the vector space R^n. The joint distribution of $\vec{x}$ is given by

$$F(\vec{x}) = F_{(X_1, X_2, \dots, X_n)}(x_1, x_2, \dots, x_n)$$

$$= F_{X_1}(x_1)\, F_{X_2}(x_2) \cdots F_{X_n}(x_n)$$

$$= F_X(x_1)\, F_X(x_2) \cdots F_X(x_n)$$

Statistics & Sampling distribution :

Let us consider a sample of size n: $x_1, x_2, \dots, x_n$ drawn from the population of a random variable X connected with a random experiment E.

Any function of sample values: $x_1, x_2, \dots, x_n$ is known as a statistic. Therefore 'a' is a statistic if it is a function of sample values : $x_1, x_2, \ldots, x_n$, i.e.,

$a = a(x_1, x_2, \ldots, x_n)$.

For example,

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i,$$

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

are statistics.

Let A is a random variable corresponding to a statistic $a = a(x_1, x_2, \ldots, x_n)$, i.e.,

$A = a(X_1, X_2, \ldots, X_n)$,

where $X_1, X_2, \ldots, X_n$ are the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively such that $X_1, X_2, \ldots, X_n$ are mutually independent and each having the same distribution as X. Therefore 'a' can be regarded as an observed value of the random variable $A = a(X_1, X_2, \cdots, X_n)$ because $a = a(x_1, x_2, \cdots, x_n)$ and each x_i is an observed value of X_i $\forall i = 1(1)n$. The distribution of A is knowns as sampling distribution.

Therefore, probability distribution of the random variable corresponding to a statistic is known as sampling distribution.

① Random variable corresponding to the sample mean is given by

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \quad \left[\text{As } \bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} \right]$$

② Random variable corresponding to the sample variance is given by

$$T^2 = \frac{1}{n}\left\{(X_1-\bar{X})^2 + (X_2-\bar{X})^2 + \cdots + (X_n-\bar{X})^2\right\}$$

$\left[\text{As sample variance } s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]$

③ $A_K = \frac{1}{n}\left(X_1^k + X_2^k + \cdots + X_n^k\right)$

is the random variable corresponding to the k th order raw moment of the sample

$$a_K = \frac{1}{n}\left(x_1^k + x_2^k + \cdots + x_n^k\right)$$

④ $M_K = \frac{1}{n}\left\{(X_1-\bar{X})^k + (X_2-\bar{X})^k + \cdots + (X_n-\bar{X})^k\right\}$

is the random variable corresponding to the k th order central moment of the sample

$$m_k = \frac{1}{n}\left\{(x_1-\bar{x})^k + (x_2-\bar{x})^k + \cdots + (x_n-\bar{x})^k\right\}$$

⑤ $G_1 = \frac{M_3}{T^3}$ corresponding to $g_1 = \frac{m_3}{s^3}$

⑥ $G_2 = \frac{M_4}{T_4} - 3$ corresponding to $g_2 = \frac{m_4}{s^4} - 3$

Theorem : $A_1 = \bar{x}$

Theorem : $M_1 = 0$

Theorem : $M_2 = T^2$

Theorem : $M_k = \sum_{i=0}^{k} (-1)^i (\bar{x})^i \binom{k}{i} A_{k-i}$

<u>Proof</u> : $M_k = \frac{1}{n}\left\{(x_1 - \bar{x})^k + (x_2 - \bar{x})^k + \cdots + (x_n - \bar{x})^k\right\}$

$$= \frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^k = \frac{1}{n}\sum_{j=1}^{n}(-\bar{x} + x_j)^k$$

$$= \frac{1}{n}\sum_{j=1}^{n}\sum_{i=0}^{k}\binom{k}{i}(-\bar{x})^i (x_j)^{k-i}$$

$$= \frac{1}{n}\sum_{j=1}^{n}\sum_{i=0}^{k}\binom{k}{i}(-1)^i \bar{x}^i (x_j)^{k-i}$$

$$= \frac{1}{n}\sum_{i=0}^{k}\sum_{j=1}^{n}\underbrace{(-1)^i\binom{k}{i}(\bar{x})^i}_{\text{free from } j}(x_j)^{k-i}$$

$$= \frac{1}{n}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(\bar{x})^i\sum_{j=1}^{n}(x_j)^{k-i}$$

$$= \sum_{i=0}^{k}(-1)^i\binom{k}{i}(\bar{x})^i A_{k-i}$$

Estimates — Good, Consistent, Unbiased:

Let α be an unknown parameter associated with the distribution function of the population of a random variable X defined on an event space S connected with a random experiment E. The functional form of the distribution function of the population of X is known but it contains the unknown parameter α, i.e., α is the one of the unknown parameters associated with the population of X. Now our task is to estimate α on the basis of the sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of X.

Consider a statistic

$$a = a(x_1, x_2, \ldots, x_n).$$

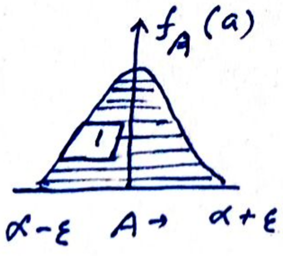
Let A be the random variable corresponding to the statistic $a = a(x_1, x_2, \dots, x_n)$, i.e.,

$A = a(X_1, X_2, \dots, X_n)$,

where $X_1, X_2, \dots, X_n$ are mutually independent random variables corresponding to the sample values $x_1, x_2, \dots, x_n$ respectively each having the same distribution as X.

① <u>Good estimate</u> : The statistic a is said to be a good estimate of α if the probability mass of the sampling distribution of the statistic a is concentrated near the point α, i.e.,

$P(|A-\alpha| < \varepsilon) \approx 1$ for small $\varepsilon > 0$



(2) <u>Consistent estimate</u> : The statistic a is said to be a consistent estimate of α if

$$\lim_{n \to \alpha} P(|A - \alpha| < \varepsilon) = 1$$

or

$$\lim_{n \to \alpha} P(|A - \alpha| \geq \varepsilon) = 0$$

or

$$A \xrightarrow[in\, p]{} \alpha \quad \text{as } n \to \infty$$

(3) <u>Unbiased estimate</u> : The statistic a is said to be an unbiased estimate of α if

$$E(A) = \alpha$$

(4) <u>Biased estimate</u> : If $E(A) \neq \alpha$, a is said to be a biased estimate of α.

(5) **Positively biased estimate:** If $E(A) - \alpha > 0$, then a is said to be positively biased estimate of α.

(6) **Negatively biased estimate:** If $E(A) - \alpha < 0$, then a is said to be negatively biased estimate of α.

(7) $E(A) - \alpha$ is known as the measure of biasedness.

Theorem: Sample mean is a consistent and unbiased estimate of population mean.

Proof: Let us consider a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from a population of a random variable X connected with a random experiment E. Then the sample mean $\bar{x}$ is given by the following equation:

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

The random variable corresponding to $\bar{x}$ is given by

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n},$$

where $X_1, X_2, \ldots, X_n$ are mutually independent random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively each having the same distribution as X

$\therefore \quad E(X_i X_j) = E(X_i)\,E(X_j) \quad \forall\, i \neq j$

$E(X_i) = E(X) = \mu \text{ (say) } \forall\, i = 1(1)n$

$Var(X_i) = Var(X) = \sigma^2 \text{ (say) } \forall\, i = 1(1)n$

$$E(\bar{X}) = E\left(\frac{1}{n}\{X_1 + X_2 + \cdots + X_n\}\right)$$

$$= \frac{1}{n} E(X_1 + X_2 + \cdots + X_n)$$

$$= \frac{1}{n}\{E(X_1) + E(X_2) + \cdots + E(X_n)\}$$

$$= \frac{1}{n}\{\mu + \mu + \cdots + \mu\} = \frac{1}{n} \cdot n\mu = \mu = E(X)$$

$$\Rightarrow \boxed{E(\bar{x}) = E(x) = \mu}$$

$\Rightarrow \bar{x}$ is an unbiased estimate of μ

$\Rightarrow$ Sample mean is an unbiased estimate of the population mean.

Again

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

$$\Rightarrow \bar{x} = \frac{1}{n} x_1 + \frac{1}{n} x_2 + \cdots + \frac{1}{n} x_n$$

$$\Rightarrow Var(\bar{x}) = \left(\frac{1}{n}\right)^2 Var(x_1) + \left(\frac{1}{n}\right)^2 Var(x_2) + \cdots + \left(\frac{1}{n}\right)^2 Var(x_n)$$

$$= \left(\frac{1}{n}\right)^2 \sigma^2 + \left(\frac{1}{n}\right)^2 \sigma^2 + \cdots + \left(\frac{1}{n}\right)^2 \sigma^2$$

$$= n \cdot \left(\frac{1}{n}\right)^2 \sigma^2 = \frac{\sigma^2}{n}$$

$$\Rightarrow \boxed{Var(\bar{x}) = \frac{\sigma^2}{n}}$$

Applying Tchebycheff's inequality on the random variable $\bar{x}$, we get

$$P(|\bar{x} - E(\bar{x})| \geq \epsilon) \leq \frac{Var(\bar{x})}{\epsilon^2}$$

for any $\epsilon > 0$

$$P(|\bar{x} - E(\bar{x})| \geq \varepsilon) \leq \frac{Var(\bar{x})}{\varepsilon^2}$$

$$\Downarrow$$

$$P(|\bar{x} - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

$$\Downarrow$$

$$0 \leq P(|\bar{x} - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

$$\Downarrow$$

$$\lim_{n\to\infty} 0 \leq \lim_{n\to\infty} P(|\bar{x} - \mu| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2}$$

$$\Downarrow$$

$$0 \leq \lim_{n\to\infty} P(|\bar{x} - \mu| \geq \varepsilon) \leq 0$$

$$\Downarrow$$

$$\lim_{n\to\infty} P(|\bar{x} - \mu| \geq \varepsilon) = 0$$

$$\Downarrow$$

$$\lim_{n\to\infty} \left[1 - P(|\bar{x} - \mu| < \varepsilon)\right] = 0$$

$$\Downarrow$$

$$\lim_{n\to\infty} P(|\bar{x} - \mu| < \varepsilon) = 1$$

$$\Downarrow$$

$$\bar{x} \xrightarrow[mp]{} \mu \text{ as } n \to \infty$$

$\Rightarrow \bar{x}$ is a consistent estimate of μ.

Therefore sample mean is a consistent and unbiased estimate of the population mean.

Theorem: The k-th order sample raw moment is a consistent and unbiased estimate of the k-th order raw moment of the population.

Proof: Let $x_1, x_2, \ldots, x_n$ be a sample of size n drawn from the population of a random variable X. Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively, where $X_1, X_2, \ldots, X_n$ are mutually independent and each having the same distribution as X. Therefore

$$E(X_i^k) = E(X^k) = \alpha_k \text{ (say)} \quad \forall\, i = 1(1)n$$

$$Var(X_i^k) = Var(X^k) \quad \forall\, i = 1(1)n$$

$$E(X_i^k X_j^k) = E(X_i^k)\, E(X_j^k) \quad \forall\, i \neq j$$

[Here we have used the following theorem of probability distribution :

If X and Y are indebendent random variables

then $E(g(X)h(Y)) = E(g(X))\,E(h(Y))$.

Taking $g(X) = X^k$, $h(Y) = Y^k$, we get

$E(X^k Y^k) = E(X^k)\,E(Y^k)$]

The k-th order sample raw moment is given by

$$a_k = \frac{x_1^k + x_2^k + \dots + x_n^k}{n}$$

The random variable corresponding to a_k is given by

$$A_k = \frac{X_1^k + X_2^k + \dots + X_n^k}{n}$$

$$\therefore\ E(A_k) = E\left(\frac{X_1^k + X_2^k + \dots + X_n^k}{n}\right)$$

$$= \frac{1}{n} E(X_1^k + X_2^k + \dots + X_n^k)$$

$$= \frac{1}{n}\left[E(X_1^k) + E(X_2^k) + \dots + E(X_n^k)\right]$$

$$= \frac{1}{n}\left[E(X^k) + E(X^k) + \dots + E(X^k)\right]$$

$$= \frac{1}{n}\left[\alpha_k + \alpha_k + \dots + \alpha_k\right]$$

$$= \frac{1}{n} \cdot n\,\alpha_k = \alpha_k = E(X^k)$$

$$\Rightarrow E(A_k) = E(X^k) = \alpha_k$$

$\Rightarrow$ a_k is an unbiased estimate of α_k

$\Rightarrow$ kth order sample raw moment is an unbiased estimate of the k-th order raw moment of the population.

$$Var(A_k) = Var\left[\frac{X_1^k + X_2^k + \cdots + X_n^k}{n}\right]$$

$$= Var\left[\frac{1}{n}X_1^k + \frac{1}{n}X_2^k + \cdots + \frac{1}{n}X_n^k\right]$$

$$= \left(\frac{1}{n}\right)^2 Var(X_1^k) + \left(\frac{1}{n}\right)^2 Var(X_2^k) + \cdots + \left(\frac{1}{n}\right)^2 Var(X_n^k)$$

$$= \left(\frac{1}{n}\right)^2 Var(X^k) + \left(\frac{1}{n}\right)^2 Var(X^k) + \cdots + \left(\frac{1}{n}\right)^2 Var(X^k)$$

$$= n\left(\frac{1}{n}\right)^2 Var(X^k) = \frac{Var(X^k)}{n}$$

Applying Tchebycheff's inequality on the random variable A_k, we get the following inequality:

$$P(|A_k - E(A_k)| \geq \varepsilon) \leq \frac{Var(A_k)}{\varepsilon^2}$$

for any fixed $\varepsilon > 0$.

$$\therefore P(|A_K - E(A_K)| \geq \varepsilon) \leq \frac{Var(A_K)}{\varepsilon^2}$$

$$\Rightarrow 0 \leq P(|A_K - E(A_K)| \geq \varepsilon) \leq \frac{Var(A_K)}{\varepsilon^2}$$

$$\Rightarrow 0 \leq P(|A_K - \alpha_K| \geq \varepsilon) \leq \frac{Var(x^k)}{n\varepsilon^2}$$

$$\Rightarrow \lim_{n\to\infty} 0 \leq \lim_{n\to\infty} P(|A_K - \alpha_K| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{Var(x^k)}{n\varepsilon^2}$$

$$\Rightarrow 0 \leq \lim_{n\to\infty} P(|A_K - \alpha_K| \geq \varepsilon) \leq 0$$

$$\Rightarrow \lim_{n\to\infty} P(|A_K - \alpha_K| \geq \varepsilon) = 0$$

$$\Rightarrow \lim_{n\to\infty} \left[1 - P(|A_K - \alpha_K| < \varepsilon)\right] = 0$$

$$\Rightarrow 1 - \lim_{n\to\infty} P(|A_K - \alpha_K| < \varepsilon) = 0$$

$$\Rightarrow \lim_{n\to\infty} P(|A_K - \alpha_K| < \varepsilon) = 1$$

$$\Rightarrow A_K \xrightarrow[mp]{} \alpha_K \text{ as } n \to \infty$$

$\Rightarrow a_K$ is a consistent estimate of α_K iff $Var(x^k)$ exists finitely

$\therefore$ a_k is a consistent and unbiased estimate of α_k iff $Var(X^k)$ exists finitely

i.e., iff $E(X^k - E(X^k))^2$ exists finitely

i.e., iff $E(X^k - \alpha_k)^2$ exists finitely

i.e., iff $E(X^{2k} - 2\alpha_k X^k + \alpha_k^2)$ exists finitely

i.e., iff $E(X^{2k}) - 2\alpha_k E(X^k) + E(\alpha_k^2)$ exists finitely

i.e., iff $\alpha_{2k} - 2\alpha_k \alpha_k + \alpha_k^2$ exists finitely

i.e., iff $\alpha_{2k} - \alpha_k^2$ exists finitely

i.e., iff α_{2k} exists finites [As existence of α_{2k} $\Rightarrow$ the existence of α_k]

<u>Theorem</u>: The k-th central moment of a sample is a consistent estimate of k-th central moment of population iff k-th central moment of population exists finitely.

<u>Proof</u>: Let $x_1, x_2, \ldots, x_n$ be a sample of size n drawn from a population of the random variable X. Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent random variables each having the same distribution as X. Then the k-th central moment of the sample is given by

$$m_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^k \quad \text{where } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

The random variable corresponding to m_k is given by

$$M_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k \quad \text{where } \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

We know that

① $M_k = \sum_{i=0}^{k} (-1)^i \binom{k}{i} A_{k-i}\, \bar{X}^i$ ② $m_k = \sum_{i=0}^{k} (-1)^i \binom{k}{i} a_{k-i}\, \bar{x}^i$

③ a_k is a consistent estimate of α_k $\forall\, k = 0, 1, 2, \ldots$

From ③, we get

(3A) $a_1 = \bar{x}$ is a consistent estimate of $\alpha_1 = E(X) = \mu$

3B a_{k-i} is a consistent estimate of α_{k-i} $\forall\, i = 0(1)k$

So, we have

④ $A_{k-i} \xrightarrow[mp]{} \alpha_{k-i}$ as $n \to \infty$ $\forall\, i = 0(1)k$

⑤ $\bar{X} \xrightarrow[mp]{} \mu$ as $n \to \infty$

$\Downarrow$

⑥ $\bar{X}^{i} \xrightarrow[mp]{} \mu^{i}$ as $n \to \infty$

From ④ and ⑥, we get

$$A_{k-i}\,\bar{X}^{i} \xrightarrow[mp]{} \alpha_{k-i}\,\mu^{i} \quad \text{as } n \to \infty$$

$\Downarrow$

$$(-1)^{i}\binom{k}{i} A_{k-i}\,\bar{X}^{i} \xrightarrow[mp]{} (-1)^{i}\binom{k}{i}\alpha_{k-i}\,\mu^{i} \quad \text{as } n \to \infty$$

$\Downarrow$

$$\sum_{i=0}^{k}(-1)^{i}\binom{k}{i} A_{k-i}\,\bar{X}^{i} \xrightarrow[mp]{} \sum_{i=0}^{k}(-1)^{i}\binom{k}{i}\alpha_{k-i}\,\mu^{i} \quad \text{as } n \to \infty$$

$\Downarrow$

$$M_k \xrightarrow[mp]{} E\left(X - E(X)\right)^{k} = E\left(X - \mu\right)^{k}$$

$$= \sum_{i=0}^{k}(-1)^{i}\binom{k}{i}\alpha_{k-i}\,\mu^{i} = \mu_k \text{ (say)}$$

$\Downarrow$

$$M_k \xrightarrow[mp]{} \mu_k \quad \text{as } n \to \infty$$

$\Rightarrow m_k$ is a consistent estimate of μ_k

$\Rightarrow$ k-th order central sample moment is a consistent estimate of k-th order central moment of the population.

Theorem: Sample variance is a consistent estimate of the population variance but it is not an unbiased estimate of population variance.

Proof: Let us consider a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of the random variable X. Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent and each having the same distribution as X. Therefore, k-th order central moment of the given sample

$$m_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^k \quad \text{where } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

we know that

m_k is consistent estimate of $\mu_k = E(X-E(X))^k$

$\Downarrow$

m_2 is a consistent estimate of $\mu_2 = E(X-E(X))^2 = \sigma^2$

$\Downarrow$

s^2(= sample variance) is a consistent estimate of the population variance σ^2, where

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

The random variable corresponding to s^2 is given by the following equation:

$$T = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \quad \text{where } \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\left[E(X_1) + E(X_2) + \cdots + E(X_n)\right]$$

$$= \frac{1}{n}\left[E(X) + E(X) + \cdots + E(X)\right] = \frac{1}{n}\cdot n\, E(X)$$

$$= E(X) = \mu$$

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum X_i\right) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)$$

$$= Var\left(\frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n\right)$$

$$= \left(\frac{1}{n}\right)^2 Var(X_1) + \left(\frac{1}{n}\right)^2 Var(X_2) + \cdots + \left(\frac{1}{n}\right)^2 Var(X_n)$$

$$\Rightarrow Var(\bar{X}) = \frac{1}{n^2} Var(x) + \frac{1}{n^2} Var(x) + \cdots + \frac{1}{n^2} Var(x)$$

$$\Rightarrow Var(\bar{X}) = \frac{1}{n^2} \cdot n \, Var(x) = \frac{\sigma^2}{n}.$$

$$T = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^{n} [x_i - \mu - (\bar{x} - \mu)]^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} [(x_i - \mu)^2 - 2(\bar{x} - \mu)(x_i - \mu) + (\bar{x} - \mu)^2]$$

$$= \frac{1}{n} \left[\sum_{i=1}^{n} (x_i - \mu)^2 - 2(\bar{x} - \mu) \sum_{i=1}^{n} (x_i - \mu) + \sum_{i=1}^{n} (\bar{x} - \mu)^2 \right]$$

$$= \frac{1}{n} \left[\sum_{i=1}^{n} (x_i - \mu)^2 - 2(\bar{x} - \mu)(n\bar{x} - n\mu) + n(\bar{x} - \mu)^2 \right]$$

$$= \frac{1}{n} \left[\sum_{i=1}^{n} (x_i - \mu)^2 - 2n(\bar{x} - \mu)^2 + n(\bar{x} - \mu)^2 \right]$$

$$= \frac{1}{n} \left[\sum_{i=1}^{n} (x_i - \mu)^2 - n(\bar{x} - \mu)^2 \right]$$

$$= \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 - (\bar{x} - \mu)^2$$

$$E(T) = \frac{1}{n} \sum_{i=1}^{n} E(x_i - \mu)^2 - E(\bar{x} - \mu)^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} Var(x_i) - Var(\bar{x}) \quad \left[\begin{array}{l} \text{as } E(x_i) = E(x) = \mu \\ \text{and } E(\bar{x}) = \mu \end{array} \right.$$

$\therefore E(T) = \frac{1}{n}[Var(x_1) + Var(x_2) + \cdots + Var(x_n)]$
$- \frac{\sigma^2}{n}$

$= \frac{1}{n}[Var(x) + Var(x) + \cdots + Var(x)] - \frac{\sigma^2}{n}$

[as $Var(x_i) = Var(x)$]

$= \frac{1}{n} n Var(x) - \frac{\sigma^2}{n} = Var(x) - \frac{\sigma^2}{n}$

$= \sigma^2 - \frac{\sigma^2}{n} = \left(1 - \frac{1}{n}\right)\sigma^2$

$\Rightarrow E(T) \neq \sigma^2$

$\Rightarrow$ s^2 is not an unbiased estimate of the population variance σ^2 but s^2 is a consistent estimate of σ^2

① Again $E(T) - \sigma^2 = -\frac{\sigma^2}{n} < 0 \Rightarrow E(T) < \sigma^2$

$\Rightarrow$ s^2 is negatively biased.

② Again $E(T) = \left(1 - \frac{1}{n}\right)\sigma^2$

$\Rightarrow E(T) = \frac{n-1}{n}\sigma^2$

$\Rightarrow \frac{n}{n-1} E(T) = \sigma^2$

$\Rightarrow E\left(\frac{n}{n-1} T\right) = \sigma^2$

$\Rightarrow E(t) = \sigma^2$, where $t = \frac{n}{n-1} T$

$\Rightarrow \frac{n}{n-1} s^2$ is an unbiased estimate of σ^2

$\Rightarrow \frac{n}{n-1} \cdot \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$ is an unbiased estimate of σ^2

$\Rightarrow \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ is an unbiased estimate of σ^2

$\Rightarrow s^2$ is an unbiased estimate of σ^2

Again,

s^2 is a consistent estimate of σ^2

$\Rightarrow T \xrightarrow[\text{in } p]{} \sigma^2$ as $n \to \infty$

$\Rightarrow T \xrightarrow[\text{in } p]{} \sigma^2$ as $n \to \infty$ and

$\checkmark \; \frac{n}{n-1} = \frac{1}{1-\frac{1}{n}} \longrightarrow 1$ as $n \to \infty$

$\Rightarrow \frac{n}{n-1} \cdot T \xrightarrow[\text{in } p]{} 1 \cdot \sigma^2$ as $n \to \infty$

$\Rightarrow t \xrightarrow[\text{in } p]{} \sigma^2$ as $n \to \infty$

$\Rightarrow s^2$ is a consistent estimate of σ^2

$\therefore s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ is known as modified sample variance and the modified sample variance is a consistent and unbiased estimate of σ^2.

Therefore, the sample variance

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \longrightarrow \underline{\underline{\text{Consistent but not unbiased}}}$$

is a consistent but not an unbiased estimate of the population variance whereas the modified sample variance

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \longrightarrow \underline{\underline{\text{Consistent and unbiased}}}$$

is a both consistent and unbiased estimate of the population variance.

It is important to note that we have frequently used the following theorem of probability:

Theorem: If $X_1, X_2, \ldots, X_n$ are mutually independent random variables then

$$Var(a_1 X_1 + a_2 X_2 + \cdots + a_n X_n)$$
$$= a_1^2 \, Var(X_1) + a_2^2 \, Var(X_2) + \cdots + a_n^2 \, Var(X_n)$$

We have also used the following theorem

Theorem: For any finite sequence of random variables $X_1, X_2, \ldots, X_n$,

$$E(a_1 X_1 + a_2 X_2 + \cdots + a_n X_n) = a_1 E(X_1) + a_2 E(X_2) + \cdots + a_n E(X_n).$$

<u>Theorem:</u> Let $X \approx N(m, \sigma)$, i.e., X is normally distributed with mean m and standard deviation σ. Then the sampling distribution of the sample mean $\bar{x}$ for a sample of size n drawn from the population of X is $N(m, \frac{\sigma}{\sqrt{n}})$.

<u>Proof:</u> Consider a sample $x_1, x_2, \ldots, x_n$ of size n drawn from the population of X. Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2 \ldots, X_n$ are mutually independent random variables each having the same distribution as X.

$\therefore X_i \approx N(m, \sigma) \quad \forall\, i = 1(1)n$

$\Rightarrow E(X_i) = m$ and $Var(X_i) = \sigma^2 \quad \forall\, i = 1(1)n$

The sample mean $\bar{x}$ and the random variable corresponding to $\bar{x}$ can be expressed as

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\therefore E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)$$

$$= E\left(\frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n\right)$$

$$= \frac{1}{n}E(X_1) + \frac{1}{n}E(X_2) + \cdots + \frac{1}{n}E(X_n)$$

$$= \frac{1}{n}m + \frac{1}{n}m + \cdots + \frac{1}{n}m$$

$$= n\left(\frac{1}{n}m\right) = n \times \frac{1}{n} \times m = m$$

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)$$

$$= Var\left(\frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n\right)$$

$$= \left(\frac{1}{n}\right)^2 Var(X_1) + \left(\frac{1}{n}\right)^2 Var(X_2) + \cdots + \left(\frac{1}{n}\right)^2 Var(X_n)$$

$$= \left(\frac{1}{n}\right)^2 \sigma^2 + \left(\frac{1}{n}\right)^2 \sigma^2 + \cdots + \left(\frac{1}{n}\right)^2 \sigma^2$$

$$= n \times \left(\frac{1}{n}\right)^2 \sigma^2 = \frac{\sigma^2}{n}$$

$$\Rightarrow \sqrt{Var(\bar{X})} = \frac{\sigma}{\sqrt{n}}$$

$\therefore \; E(\bar{x}) = m$ and $\sqrt{Var(\bar{x})} = \frac{\sigma}{\sqrt{n}}$

Again we know the following theorem of probability:

<u>Theorem:</u> If $X_i \approx N(m_i, \sigma_i) = N(m_i, \sigma_i)$ and $X_1, X_2, \ldots, X_n$ are mutually independent random variables, then for any set of real constants $a_1, a_2, \ldots, a_n$; the random variable $a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$ is $N(m, \sigma)$ where

$$m = a_1 m_1 + a_2 m_2 + \cdots + a_n m_n$$

and

$$\sigma^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \cdots + a_n \sigma_n^2$$

So, according to this theorem, $\bar{X} \approx N(m, \frac{\sigma}{\sqrt{n}})$

$$\Rightarrow \frac{\bar{x} - m}{(\sigma / \sqrt{n})} \sim N(0, 1)$$

$\Rightarrow$ The sampling distribution of $\bar{x}$ is $N(m, \frac{\sigma}{\sqrt{n}})$

or

the sampling distribution of the statistic

$\frac{\bar{x} - m}{\sigma / \sqrt{n}} = \frac{\sqrt{n}}{\sigma}(\bar{x} - m)$ is standard normal.

<u>Theorem</u>: Let $X \approx N(m, \sigma)$. Then the statistic $\frac{ns^2}{\sigma^2}$ has a χ^2- distribution with $\nu = n-1$ degrees of freedom where

① s^2 is the sample variance for a sample of size n drawn from the population of X

② sample mean and sample variance are independet.

<u>Proof</u>: Let $x_1, x_2, \dots, x_n$ be a sample of size n drawn from the population $X \sim N(m, \sigma)$. Let $X_1, X_2, \dots, X_n$ be the randam variables corresponding to the sample values $x_1, x_2, \dots, x_n$ respectively. Therefore, $X_1, X_2, \dots, X_n$ are mutually independent random variables and $X_i \sim N(m, \sigma)$ $\forall\, i = 1(1)n$. The sample variance s^2 is given by

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \quad \text{where } \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

The random variable T corresponding to s^2 is given by

$$T = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 \quad \text{where } \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

According to the condition of the problem:
T and $\bar{X}$ are independent random variables.

Now the random variable corresponding to the statistic $\frac{ns^2}{\sigma^2}$ is given by

$$Y = \frac{nT}{\sigma^2} = \frac{n}{\sigma^2} \cdot \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$= \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$= \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n} \left(\frac{X_i - m - (\bar{X} - m)}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n} \left[\left(\frac{X_i - m}{\sigma}\right)^2 - 2 \cdot \frac{X_i - m}{\sigma} \cdot \frac{\bar{X} - m}{\sigma} + \left(\frac{\bar{X} - m}{\sigma}\right)^2\right]$$

$$= \sum_{i=1}^{n} \left(\frac{X_i - m}{\sigma}\right)^2 - 2 \cdot \frac{\bar{X} - m}{\sigma} \sum_{i=1}^{n} \frac{X_i - m}{\sigma} + \sum_{i=1}^{n} \left(\frac{\bar{X} - m}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n} \left(\frac{X_i - m}{\sigma}\right)^2 - 2\, \frac{\bar{X} - m}{\sigma} \cdot \frac{n(\bar{X} - m)}{\sigma} + n\left(\frac{\bar{X} - m}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n} \left(\frac{X_i - m}{\sigma}\right)^2 - n\left(\frac{\bar{X} - m}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n} \left(\frac{X_i - m}{\sigma}\right)^2 - \left\{\frac{\bar{X} - m}{(\sigma/\sqrt{n})}\right\}^2$$

$$\therefore\ Y = \frac{nT}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{x_i - m}{\sigma}\right)^2 - \left[\frac{\bar{x} - m}{\sigma/\sqrt{n}}\right]^2$$

$$= \sum_{i=1}^{n} Y_i^2 - \left[\frac{\bar{x} - m}{\sigma/\sqrt{n}}\right]^2 \quad \text{where } Y_i = \frac{x_i - m}{\sigma}$$

Now $\dfrac{\bar{x} - m}{(\sigma/\sqrt{n})} = \dfrac{\sqrt{n}}{\sigma}\left[\bar{x} - m\right] = \dfrac{\sqrt{n}}{\sigma}\left[\dfrac{1}{n}\sum_{i=1}^{n} x_i - m\right]$

$$= \frac{\sqrt{n}}{\sigma . n}\left(\sum x_i - mn\right)$$

$$= \frac{1}{\sqrt{n}} \cdot \frac{1}{\sigma} \sum_{i=1}^{n} (x_i - m)$$

$$= \sum_{i=1}^{n} \frac{1}{\sqrt{n}} \frac{x_i - m}{\sigma} = \sum_{i=1}^{n} \frac{1}{\sqrt{n}} Y_i$$

$$= \sum_{i=1}^{n} a_{1i} Y_i$$

where $a_{1i} = \frac{1}{\sqrt{n}} \quad \forall\ i = 1(1)n$

$$a_{11}^2 + a_{12}^2 + \cdots + a_{1n}^2 = \frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n} = n\left(\frac{1}{n}\right) = 1$$

$$\therefore\ Y = \sum_{i=1}^{n} Y_i^2 - \left(\sum_{i=1}^{n} a_{1i} Y_i\right)^2$$

where $Y_i \sim N(0,1)\ \forall\ i = 1(1)n$ as $x_i \sim N(m, \sigma)$

$$\Downarrow$$

$$\frac{x_i - m}{\sigma} \sim N(0,1)$$

$$\Downarrow$$

$$Y_i \sim N(0,1)$$

and $\sum_{i=1}^{n} a_{1i} Y_i$ is a linear combination satisfying orthogonality condition $\sum_{i=1}^{n} a_{1i}^2 = 1$

therefore y is a χ^2- variate with $(n-1)$ degrees of freedom $\Rightarrow \frac{nT}{\sigma^2}$ is a χ^2- variate with $(n-1)$ degrees of freedom $\Rightarrow$ sampling distribution of $\frac{ns^2}{\sigma^2}$ is a χ^2 distribution with $(n-1)$ degrees of freedom.

<u>Theorem</u>: The sampling distribution of the statistic $t = \frac{\sqrt{n}(\bar{x}-m)}{s}$ is a t- distribution of $(n-1)$ degrees of freedom, where $\bar{x}$ is the sample mean and s is the modified sample variance.

<u>Proof</u>: Let $x_1, x_2, \ldots, x_n$ be a sample of size n drawn from the population of the random variable X, where $X \sim N(m, \sigma)$. Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent and each $X_i \sim N(m, \sigma)$ $\forall i = 1(1)n$.

We know that

① the sampling distribution of $u = \frac{\bar{x} - m}{(\sigma/\sqrt{n})}$ is $N(0,1)$

② the sampling distribution of $\frac{ns^2}{\sigma^2}$ is a

χ^2- distribution with $(n-1)$ degrees of freedom.

$$\therefore \quad U = \frac{\bar{X} - m}{\sigma}, \quad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$Y = \frac{nT}{\sigma^2}, \quad T = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

So, ③ U is the random variable corresponding to

the statistic $u = \frac{\bar{x} - m}{(\sigma/\sqrt{n})} \Rightarrow \boxed{U \sim N(0,1)}$

④ T is the random variable corresponding to

the statistic $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$

⑤ Y is the random variable corresponding to

the statistic $\frac{ns^2}{\sigma^2} \Rightarrow Y$ is χ^2- distributed

with $(n-1)$ degrees of freedom

Therefore, $\dfrac{U}{\sqrt{Y/(n-1)}}$ is t - distributed with

$(n-1)$ degrees of freeodom.

Observed value of the random variable

$$\frac{U}{\sqrt{Y/(n-1)}} = \frac{(\bar{x}-m)/(\sigma/\sqrt{n})}{\sqrt{\frac{1}{n-1}\cdot\frac{ns^2}{\sigma^2}}}$$

$$= \frac{(\bar{x}-m)/(\sigma/\sqrt{n})}{\sigma\sqrt{\frac{n}{n-1}s^2}}$$

$$= \frac{\sqrt{n}(\bar{x}-m)}{\sqrt{\frac{n}{n-1}\cdot\frac{1}{n}\cdot\sum(x_i-\bar{x})^2}} \qquad s^2 = \frac{1}{n}\sum(x_i-\bar{x})^2.$$

$$= \frac{\sqrt{n}(\bar{x}-m)}{\sqrt{\frac{1}{n-1}\sum(x_i-\bar{x})^2}}$$

$$= \frac{\sqrt{n}(\bar{x}-m)}{\sqrt{s^2}} \qquad s^2 = \frac{1}{n-1}\sum(x_i-\bar{x})^2$$

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

$\therefore$ The random variable corresponding to the statistic $\frac{\sqrt{n}(\bar{x}-m)}{s}$ is $\frac{U}{\sqrt{Y/(n-1)}}$

$\frac{U}{\sqrt{Y/(n-1)}}$ is t distributed with $(n-1)$ df.

$\Rightarrow$ sampling distribution of $\frac{\sqrt{n}(\bar{x}-m)}{s}$ is a t- distribution with $(n-1)$ df.

Problem 1: Let $x_1, x_2, \ldots, x_6$ be a sample of size 6 drawn from a normal $N(m, \sigma)$ population. Determine the constant c such that the statistic $c[(x_1 - x_2)^2 + (x_3 - x_4)^2 + (x_5 - x_6)^2]$ is an unbiased estimate of σ^2.

Answer: Let $X_1, X_2, \ldots, X_6$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_6$ respectively. Therefore, $X_1, X_2, \ldots, X_6$ are mutually independent each having the same distribution as the parent random variable $X = N(m, \sigma)$.

$$\therefore \quad E(X_i) = E(X) = m, \quad Var(X_i) = Var(X) = \sigma^2$$

$$E(X_i X_j) = E(X_i) E(X_j) = m^2 \quad \forall\, i \neq j$$

The random variable corresponding to the given statistic $[(x_1 - x_2)^2 + (x_3 - x_4)^2 + (x_5 - x_6)^2]$ is

$$A = c[(X_1 - X_2)^2 + (X_3 - X_4)^2 + (X_5 - X_6)^2]$$

We want to find the value c for which

$$E(A) = \sigma^2 \qquad (1)$$

$$A = c\left[(x_1 - x_2)^2 + (x_3 - x_4)^2 + (x_5 - x_6)^2\right]$$

$$= c\left[\{x_1 - m - (x_2 - m)\}^2 + \{x_3 - m - (x_4 - m)\}^2 + \{x_5 - m - (x_6 - m)\}^2\right]$$

$$= c\left[(x_1 - m)^2 + (x_2 - m)^2 - 2(x_1 - m)(x_2 - m) + (x_3 - m)^2 + (x_4 - m)^2 - 2(x_3 - m)(x_4 - m) + (x_5 - m)^2 + (x_6 - m)^2 - 2(x_5 - m)(x_6 - m)\right]$$

$$= c\left[\sum_{i=1}^{6}(x_i - m)^2 - 2(x_1 - m)(x_2 - m) - 2(x_3 - m)(x_4 - m) - 2(x_5 - m)(x_6 - m)\right] \quad \text{——— (2)}$$

✓ Now $E(x_i - m)^2 = E(x_i - E(x_i))^2 = Var(x_i) = Var(x) = \sigma^2$

$$\Rightarrow E(x_i - m)^2 = \sigma^2 \quad \forall\, i = 1, 2, \ldots, 6 \quad \text{——— (3)}$$

✓ $E(x_1 - m)(x_2 - m) = E(x_1 x_2 - m x_2 - m x_1 + m^2)$

$= E(x_1 x_2) - m\,E(x_2) - m\,E(x_1) + E(m^2)$

$= E(x_1)\,E(x_2) - m.m - m.m + m^2$

$= m.m - m.m - m.m + m^2 = 0$

$$\left.\begin{array}{l} \Rightarrow E(x_1 - m)(x_2 - m) = 0 \\ \text{IIly } E(x_3 - m)(x_4 - m) = 0 \\ \text{and } E(x_5 - m)(x_6 - m) = 0 \end{array}\right\} \quad \text{——— (4)}$$

From ② we get

$$E(A) = c\left[\sum_{i=1}^{6} E(x_i - m)^2 - 2E(x_1-m)(x_2-m)\right.$$

$$\left. - 2E(x_3-m)(x_4-m) - 2E(x_5-m)(x_6-m)\right]$$

$$= c\left[\sigma^2+\sigma^2+\sigma^2+\sigma^2+\sigma^2+\sigma^2 - 2\times 0 - 2\times 0 - 2\times 0\right]$$

$$= 6c\sigma^2$$

$$\Rightarrow E(A) = 6c\sigma^2$$

$$\Rightarrow \sigma^2 = 6c\sigma^2 \quad [\text{using } ①]$$

$$\Rightarrow 6c = 1$$

$$\Rightarrow \boxed{c = \frac{1}{6}}$$

<u>Problem 2</u>: Let $x_1, x_2, \dots, x_n$ be a sample of size n drawn from the population of a random variable $X = N(m, \sigma)$. Find the relation between a, b, m, σ such that the statistic

$$\sum_{i=2}^{n} (a x_i + b x_{i-1})^2$$ is an unbiased estimate of σ^2.

Answer: Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the samble values $x_1, x_2, \ldots, x_n$. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent random variables such that each X_i has the same probability distribution as X.

$\therefore X_i = N(m, \sigma) \quad \forall i = 1(1)n$

$\Rightarrow E(X_i) = m$ and $Var(X_i) = \sigma^2 \quad \forall i = 1(1)n$

$E(X_i - m)(X_j - m) = E(X_i X_j - m X_j - m X_i + m^2)$

$= E(X_i X_j) - m E(X_j) - m E(X_i) + E(m^2)$

$= E(X_i X_j) - m.m - m.m + m^2$

$= E(X_i X_j) - m^2$

$= E(X_i) E(X_j) - m^2 \quad \forall i \neq j$

$= m.m - m^2 = 0$

$\therefore E(X_i - m)(X_j - m) = 0 \quad \forall i \neq j$

The random variable corresponding to the given statistic can be written as

$$A = \sum_{i=2}^{n} (a x_i + b x_{i-1})^2$$

$$= \sum_{i=2}^{n} \{a(x_i - m) + b(x_{i-1} - m) + am + bm\}^2$$

$$= \sum_{i=2}^{n} \Big[a^2 (x_i - m)^2 + b^2 (x_{i-1} - m)^2 + (am + bm)^2$$
$$+ 2ab(x_i - m)(x_{i-1} - m)$$
$$+ 2a(am + bm)(x_i - m)$$
$$+ 2b(am + bm)(x_{i-1} - m)\Big]$$

$$E(A) = \sum_{i=2}^{n} \Big[a^2 E(x_i - m)^2 + b^2 E(x_{i-1} - m) + E(am + bm)^2$$
$$+ 2ab\, E(x_i - m)(x_i$$
$$+ 2a(am + bm)\, E(x_i - m)$$
$$+ 2b(am + bm)\, E(x_{i-1} - m)\Big]$$

$$= \sum_{i=2}^{n} \Big[a^2 Var(x_i) + b^2 Var(x_{i-1}) + (am + bm)^2$$
$$+ 2ab \cdot 0 + 2a(am + bm) \cdot 0$$
$$+ 2b(am + bm) \cdot 0\Big]$$

$$= \sum_{i=2}^{n} \left[a^2 \sigma^2 + b^2 \sigma^2 + m^2 (a + b)^2\right]$$

$$= (n-1)\left[a^2 \sigma^2 + b^2 \sigma^2 + m^2 (a + b)^2\right]$$

The given statistic is an unbiased estimate of σ^2

if $E(A) = \sigma^2$

i.e., if $(n-1)\left[a^2\sigma^2 + b^2\sigma^2 + m^2(a+b)^2\right] = \sigma^2$

i.e., if $a^2\sigma^2 + b^2\sigma^2 + m^2(a+b)^2 = \frac{\sigma^2}{n-1}$ ——— (1)

i.e., if $m^2(a+b)^2 = \sigma^2\left[\frac{1}{n-1} - a^2 - b^2\right]$

i.e., if $\frac{m^2}{\sigma^2} = \frac{\frac{1}{n-1} - a^2 - b^2}{(a+b)^2}$

As $\frac{m^2}{\sigma^2} > 0$, $\frac{1}{n-1} - a^2 - b^2 > 0$

$$\Downarrow$$

$$\frac{1}{n-1} > a^2 + b^2$$

(1) the required relation is

$$\frac{m^2}{\sigma^2} = \frac{\frac{1}{n-1} - a^2 - b^2}{(a+b)^2}$$

along with the following restriction on a, b:

✓ $a^2 + b^2 < \frac{1}{n-1}$

✓ $a + b \neq 0$

(2) If $a+b=0$, i.e., if $b = -a$ then from (1), we get

$a^2\sigma^2 + (-a)^2\sigma^2 = \frac{\sigma^2}{n-1}$

$\Rightarrow 2a^2\sigma^2 = \frac{\sigma^2}{n-1}$

$\Rightarrow a^2 = \frac{1}{2(n-1)}$

$\Rightarrow a = \pm\sqrt{\frac{1}{2(n-1)}}$

$b = \mp\sqrt{\frac{1}{2(n-1)}}$

Problem 3 : Let $x_1, x_2, \ldots, x_n$ be a sample of size n drawn from a population of a random variable X. If $Var(X)$ exists finitely, then determine the value of d such that the statistic $\sum_{i=1}^{n} d(x_i - \bar{x})^2$ with $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is an unbiased estimate of the population variance $\sigma^2 = Var(X)$.

Answer: $Var(X)$ exists finitely $\Rightarrow$ $E(X)$ exists finitely

Let $m = E(X)$ and $\sigma^2 = Var(X)$. Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, the random variables $X_1, X_2, \ldots, X_n$ are mutually independent and the distribution of X_i is same as X $\forall i = 1(1)n$.

$\therefore$ $E(X_i) = m$ and $Var(X_i) = \sigma^2$ $\forall i = 1(1)n$

$E(X_i X_j) = E(X_i)\,E(X_j)$ $\forall i \neq j$.

The random variable corresponding to the given statistic can be presented as follows:

$$A = \sum_{i=1}^{n} d\,(x_i - \bar{x})^2 \quad \text{with } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Now

$$A = d\sum_{i=1}^{n} (x_i - \bar{x})^2 = d\sum_{i=1}^{n}\left[x_i - m - (\bar{x} - m)\right]^2$$

$$= d\sum_{i=1}^{n}\left[(x_i - m)^2 - 2(\bar{x}-m)(x_i - m) + (\bar{x}-m)^2\right]$$

$$= d\left[\sum_{i=1}^{n}(x_i - m)^2 - 2(\bar{x}-m)\sum_{i=1}^{n}(x_i - m) + \sum_{i=1}^{n}(\bar{x}-m)^2\right]$$

$$= d\left[\sum_{i=1}^{n}(x_i - m)^2 - 2(\bar{x}-m)(n\bar{x} - nm) + n(\bar{x}-m)^2\right]$$

$$= d\left[\sum_{i=1}^{n}(x_i - m)^2 - n(\bar{x}-m)^2\right]$$

$$\therefore E(A) = d\left[\sum_{i=1}^{n} E(x_i - m)^2 - n\,E(\bar{x}-m)^2\right]$$

$$= d\left[\sum_{i=1}^{n} Var(x_i) - n\,E(\bar{x}-m)^2\right]$$

$$= d\left[\sum_{i=1}^{n} \sigma^2 - n\,E(\bar{x}-m)^2\right]$$

$$= d\left[n\sigma^2 - n\,E(\bar{x}-m)^2\right]$$

Again $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

$$\Rightarrow E(\bar{x}) = \frac{1}{n} \sum_{i=1}^{n} E(x_i) = \frac{1}{n} \sum_{i=1}^{n} m = \frac{1}{n} mn = m$$

$$\Rightarrow E(\bar{x} - m)^2 = E(\bar{x} - E(\bar{x}))^2 = Var(\bar{x})$$

$$= Var \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$= Var\left(\frac{1}{n} x_1 + \frac{1}{n} x_2 + \cdots + \frac{1}{n} x_n\right)$$

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

$$= \left(\frac{1}{n}\right)^2 \sigma^2 + \left(\frac{1}{n}\right)^2 \sigma^2 + \cdots + \left(\frac{1}{n}\right)^2 \sigma^2$$

$$= n \left(\frac{1}{n}\right)^2 \sigma^2 = \frac{1}{n} \sigma^2$$

$$\Rightarrow n E(\bar{x} - m)^2 = \sigma^2$$

$$\therefore E(A) = d\left[n\sigma^2 - n E(\bar{x} - m)^2\right] = d\left[n\sigma^2 - \sigma^2\right]$$

$$\Rightarrow E(A) = d(n-1)\sigma^2$$

Now the given statistic is an unbiased estimate of σ^2 if $E(A) = \sigma^2$

i.e., if $d(n-1)\sigma^2 = \sigma^2$

i.e., if $d = \frac{1}{n-1}$

$\Rightarrow \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ is an unbiased estimate of σ^2

$\Rightarrow$ Modified sample variance is an unbiased estimate of the population variance.

<u>Problem 4</u>: Let $x_1, x_2, \ldots, x_n$ be a sample of size n drawn from a population of the random variable X, where X is a Poisson variate with parameter μ. Find the constants a and b such that the statistic

$$\sum_{i=2}^{n} x_{i-1}(a x_{i-1} + b x_i)$$ is an unbiased estimate of population variance whereas the statistic

$$\sum_{i=2}^{n} (a x_{i-1} + b x_i)$$ is an unbiased estimate of the population mean.

<u>Answer:</u> Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent random variables and the distribution of X_i is same as X for all $i = 1(1)n$.

$$\left.\begin{aligned} \therefore\ E(X_i) &= E(X) = \mu \quad \forall\ i = 1(1)n \\ Var(X_i) &= Var(X) = \mu \quad \forall\ i = 1(1)n \\ E(X_i X_j) &= E(X_i)E(X_j) = \mu \cdot \mu = \mu^2 \quad \forall\ i \neq j \end{aligned}\right\} \text{——— (1)}$$

Now the random variables corresponding to the statistic

$\sum_{i=2}^{n} (a x_{i-1} + b x_i)$ and $\sum_{i=2}^{n} x_{i-1} (a x_{i-1} + b x_i)$ are

respectively given by

$$T = \sum_{i=2}^{n} (a X_{i-1} + b X_i) \qquad (2)$$

$$U = \sum_{i=2}^{n} X_{i-1} (a X_{i-1} + b X_i) \qquad (3)$$

According to the conditions,

$$E(T) = E(X) = \mu \qquad (4)$$

$$E(U) = Var(X) = \mu \qquad (5)$$

From (4), we get

$$E(T) = \mu \Rightarrow E\left[\sum_{i=2}^{n} (a X_{i-1} + b X_i)\right] = \mu$$

$$\Rightarrow \sum_{i=2}^{n} \left[a E(X_{i-1}) + b E(X_i)\right] = \mu$$

$$\Rightarrow \sum_{i=2}^{n} (a\mu + b\mu) = \mu$$

$$\Rightarrow (a\mu + b\mu)(n-1) = \mu$$

$$\Rightarrow (a+b)\mu(n-1) = \mu$$

$$\Rightarrow \boxed{a + b = \frac{1}{n-1}} \qquad (6)$$

From 5 we get

$$E(U) = \mu \Rightarrow E\left[\sum_{i=2}^{n} x_{i-1}(a x_{i-1} + b x_i)\right] = \mu$$

$$\Rightarrow E\left[\sum_{i=2}^{n} (x_{i-1} - \mu + \mu)(a x_{i-1} + b x_i)\right] = \mu$$

$$\Rightarrow E\left[\sum_{i=2}^{n} (x_{i-1} - \mu)(a x_{i-1} + b x_i) + \sum_{i=2}^{n} \mu (a x_{i-1} + b x_i)\right] = \mu$$

$$\Rightarrow \sum_{i=2}^{n} E[(x_{i-1} - \mu)(a x_{i-1} + b x_i)] + \sum_{i=2}^{n} \mu [a E(x_{i-1}) + b E(x_i)] = \mu$$

$$\Rightarrow \sum_{i=2}^{n} E[(x_{i-1} - \mu)(a x_{i-1} + b x_i)] + \sum_{i=2}^{n} \mu [a\mu + b\mu] = \mu$$

$$\Rightarrow \sum_{i=2}^{n} E[(x_{i-1} - \mu)(a x_{i-1} + b x_i)] + (n-1)(a+b)\mu^2 = \mu$$

$$\Rightarrow \sum_{i=2}^{n} E[(x_{i-1} - \mu)(a x_{i-1} + b x_i) + (n-1)\frac{1}{n-1}\mu^2 = \mu$$

[using 6]

$$\Rightarrow \sum_{i=2}^{n} E[(x_{i-1} - \mu)(a x_{i-1} + b x_i)] + \mu^2 = \mu$$

$$\Rightarrow \sum_{i=2}^{n} E[(x_{i-1} - \mu)(a x_{i-1} + b x_i)] = \mu - \mu^2 \quad \text{——— } (7)$$

Now $E[(x_{i-1} - \mu)(a x_{i-1} + b x_i)]$

$$= E[(x_{i-1} - \mu)\{a(x_{i-1} - \mu + \mu) + b(x_i - \mu + \mu)\}]$$

$$= E[(x_{i-1} - \mu)\{a(x_{i-1} - \mu) + b(x_i - \mu) + a\mu + b\mu\}]$$

$$= a E(x_{i-1} - \mu)^2 + b E(x_{i-1} - \mu)(x_i - \mu) + (a\mu + b\mu) E(x_{i-1} - \mu)$$

$$= a\,Var(X_{i-1}) + b\,E\{X_i X_{i-1} - \mu X_i - \mu X_{i-1} + \mu^2\}$$
$$+ (a\mu + b\mu)[E(X_{i-1}) - E(\mu)]$$

$$= a\mu + b[E(X_i X_{i-1}) - \mu E(X_i) - \mu E(X_{i-1}) + E(\mu^2)]$$
$$+ (a\mu + b\mu)[\mu - \mu]$$

$$= a\mu + b[E(X_i)E(X_{i-1}) - \mu\cdot\mu - \mu\cdot\mu + \mu^2]$$

$$= a\mu + b[\mu\cdot\mu - \mu^2] = a\mu \qquad (8)$$

From (7) & (8), we get

$$\sum_{i=2}^{n} a\mu = \mu - \mu^2$$

$$\Rightarrow (n-1)a\mu = \mu(1-\mu)$$

$$\Rightarrow a = \frac{1-\mu}{n-1}$$

$$\therefore b = \frac{1}{n-1} - \frac{1-\mu}{n-1} = \frac{\mu}{n-1}$$

$$\therefore \boxed{a = \frac{1-\mu}{n-1} \quad \& \quad b = \frac{\mu}{n-1}}$$

Problem 5: Let $x_1, x_2, \ldots, x_n$ be a sample of size n drawn from a population of the random variable X, where the probability mass function of the random variable X is given by

$$P(X=i) = p(1-p)^{i-1}, \quad i = 1, 2, \ldots \quad \text{with } 0<p<1.$$

Find the constants a and b such that the statistic $\sum_{i=2}^{n} b\, x_{i-1}(a x_i - x_{i-1})$ is an unbiased estimator of $\frac{1}{p}$.

Answer:

$$E(X) = \sum_{i=1}^{\infty} i\, P(X=i) = \sum_{i=1}^{\infty} i\, p\, q^{i-1}, \quad q = 1-p$$

$$\Rightarrow \frac{E(X)}{p} = \sum_{i=1}^{\infty} i\, q^{i-1} = 1 + 2q + 3q^2 + \cdots \quad ①$$

$$\Rightarrow q\,\frac{E(X)}{p} = \quad q + 2q^2 + \cdots \quad ②$$

$$\therefore ① - ② \Rightarrow (1-q)\frac{E(X)}{p} = 1 + q + q^2 + \cdots = \frac{1}{1-q}$$

$$\Rightarrow E(X) = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

$$\Rightarrow \boxed{E(X) = \frac{1}{p}}$$

$$S = E(X(X-1)) = \sum_{i=1}^{\infty} i(i-1)\,P(X=i) = \sum_{i=1}^{\infty} i(i-1)\,p\,q^{i-1}$$

$$\Rightarrow S = 2pq + 6pq^2 + 12pq^3 + \cdots$$

$$\Rightarrow \frac{S}{2pq} = 1 + 3q + 6q^2 + \cdots$$

$$\Rightarrow q\,\frac{S}{2pq} = \quad q + 3q^2 + \cdots$$

$$\therefore (1-q)\frac{S}{2pq} = 1 + 2q + 3q^2 + \cdots$$

$$\Rightarrow q(1-q)\frac{S}{2pq} = \quad q + 2q^2 + \cdots$$

$$\therefore (1-q)\frac{S}{2pq}(1-q) = 1 + q + q^2 + \cdots = \frac{1}{1-q}$$

$$\Rightarrow (1-q)^2\frac{S}{2pq} = \frac{1}{1-q}$$

$$\Rightarrow S = \frac{2pq}{(1-q)^3} = \frac{2pq}{p^3} = \frac{2q}{p^2}$$

$$\therefore E(X(X-1)) = \frac{2q}{p^2}$$

$$\Rightarrow E(X^2) - E(X) = \frac{2q}{p^2}$$

$$\Rightarrow E(X^2) = \frac{2q}{p^2} + E(X) = \frac{2q}{p^2} + \frac{1}{p} = \frac{2q+p}{p^2}$$

$$= \frac{2(1-p)+p}{p^2} = \frac{2-p}{p^2}$$

$$\therefore Var(X) = E(X - E(X))^2 = E(X^2) - (E(X))^2$$

$$= \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{2-p-1}{p^2} = \frac{1-p}{p^2} = \frac{q}{p^2}$$

$$\Rightarrow \boxed{Var(X) = \frac{q}{p^2}}$$

Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent and the distribution function of each X_i is same as X for all $i = 1(1)n$.

$$\therefore \left.\begin{aligned} E(X_i) &= E(X) = \frac{1}{p} \quad \forall\, i = 1(1)n \\ Var(X_i) &= Var(X) = \frac{q}{p^2} \quad \forall\, i = 1(1)n \\ E(X_i X_j) &= E(X_i)\,E(X_j) = \frac{1}{p^2} \quad \forall\, i \neq j \end{aligned}\right\} \text{—— } (3)$$

Let us write $\boxed{E(X) = \mu = \frac{1}{p}}$

The given statistic y can be written as

$$y = \sum_{i=2}^{n} b\, x_{i-1}\,(a x_i - x_{i-1})$$

$$= b \sum_{i=2}^{n} (x_{i-1} - \mu + \mu)(a x_i - x_{i-1})$$

$$= b \sum_{i=2}^{n} (x_{i-1} - \mu)(a x_i - x_{i-1}) + b\mu \sum_{i=2}^{n} (a x_i - x_{i-1})$$

$$= b \sum_{i=2}^{n} (x_{i-1} - \mu)\{a(x_i - \mu + \mu) - (x_{i-1} - \mu + \mu)\}$$
$$+ b\mu \sum_{i=2}^{n} (a x_i - x_{i-1})$$

$$= b \sum_{i=2}^{n} (x_{i-1} - \mu)\{a(x_i - \mu) - (x_{i-1} - \mu) + a\mu - \mu\}$$
$$+ b\mu \sum_{i=2}^{n} (a x_i - x_{i-1})$$

$$y = b\sum_{i=2}^{n}(x_{i-1}-\mu)\{a(x_i-\mu)-(x_{i-1}-\mu)+(a\mu-\mu)\}$$
$$+ b\mu\sum_{i=2}^{n}(ax_i - x_{i-1})$$

$$= ab\sum_{i=2}^{n}(x_{i-1}-\mu)(x_i-\mu) - b\sum_{i=2}^{n}(x_{i-1}-\mu)^2$$
$$+ b(a\mu-\mu)\sum_{i=2}^{n}(x_{i-1}-\mu) + b\mu\sum_{i=2}^{n}(ax_i - x_{i-1})$$

$\therefore$ the random variable corresponding to y is given by

$$Y = ab\sum_{i=2}^{n}(X_{i-1}-\mu)(X_i-\mu) - b\sum_{i=2}^{n}(X_{i-1}-\mu)^2$$
$$+ b\mu(a-1)\sum_{i=2}^{n}(X_{i-1}-\mu) + b\mu\sum_{i=2}^{n}(aX_i - X_{i-1})$$

Therefore the given statistic is an unbiased estimate of population mean $E(X) = \frac{1}{p}$ if

$$E(Y) = \frac{1}{p}$$

i.e., if $ab\sum_{i=2}^{n}E(X_{i-1}-\mu)(X_i-\mu) - b\sum_{i=2}^{n}E(X_{i-1}-\mu)^2$

$$+ b\mu(a-1)\sum_{i=2}^{n}E(X_{i-1}-\mu) + b\mu\sum_{i=2}^{n}E(aX_i - X_{i-1})$$

$$= \frac{1}{p} \qquad \text{——————} \quad (4)$$

Now $E(x_{i-1}-\mu)(x_i-\mu) = E(x_{i-1}x_i - \mu x_i - \mu x_{i-1} + \mu^2)$

$= E(x_{i-1}x_i) - \mu E(x_i) - \mu E(x_{i-1}) + \mu^2$

$= E(x_{i-1})E(x_i) - \mu\cdot\mu - \mu\cdot\mu + \mu^2$

$= \mu\cdot\mu - \mu\cdot\mu - \mu\cdot\mu + \mu^2 = 0$

$E(x_{i-1}-\mu)^2 = Var(x_{i-1}) = Var(x) = \frac{q}{p^2}$

$E(x_{i-1}-\mu) = E(x_{i-1}) - E(\mu) = \mu - \mu = 0$

$E(ax_i - x_{i-1}) = E(ax_i) - E(x_{i-1}) = aE(x_i) - E(x_{i-1})$

$= a\mu - \mu$

$\therefore\ ab\cdot 0 - b\sum_{i=2}^{n}\frac{q}{p^2} + l\mu(a-1)\cdot 0 + b\mu\sum_{i=2}^{n}(a\mu-\mu) = \frac{1}{p}$

$\Rightarrow -b(n-1)\frac{q}{p^2} + b\mu(n-1)(a\mu-\mu) = \frac{1}{p}$

$\Rightarrow -b(n-1)\frac{1-p}{p^2} + b(n-1)(a-1)\mu^2 = \frac{1}{p}$

$\Rightarrow -b(n-1)\frac{1-p}{p^2} + b(n-1)(a-1)\cdot\frac{1}{p^2} = \frac{1}{p}$

$\Rightarrow -b(n-1)(1-p) + b(n-1)(a-1) = p$

$\Rightarrow -(1-p) + a - 1 = \frac{p}{b(n-1)}$

$\Rightarrow -1 + p + a - 1 = \frac{p}{b(n-1)}$

$\Rightarrow a - 2 = p\left[\frac{1}{b(n-1)} - 1\right]$ ———————— (5)

Equation 5 holds good for any p, $0<p<1$ if

each side of 5 is equal to zero, i.e., if
$a-2=0$ and $\frac{1}{b(n-1)}-1=0$, i.e., if $a=2$ and $b=\frac{1}{n-1}$

Therefore, the given statistic is an unbiased estimator of $1/p$ if $a=2$ and $b=\frac{1}{n-1}$

Estimation of parameters

Maximum likelihood method: Let us suppose that the distribution function of the population has a known functional form but it contains a number of unknown parameters. We want to estimate the values of the parameters on the basis of a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of the random variable X.

Likelihood Function: Let X be the parent random variable whose distribution function contains a number of unknown parameters $\theta_1, \theta_2, \ldots, \theta_k$. Consider a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of the random variable X. Then the the likelihood function is denoted by

$L \equiv L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k)$ and defined by

$$L = \begin{cases} P(X_1 = x_1)\, P(X_2 = x_2) \cdots P(X_n = x_n) & \text{for discrete case,} \\ f_{X_1}(x_1; \theta_1, \ldots, \theta_k)\, f_{X_2}(x_2; \theta_1, \ldots, \theta_k) \cdots f_{X_n}(x_n; \theta_1, \ldots, \theta_k) & \text{for continuous case,} \end{cases}$$

where

① $P(X=x)$ is the probability mass function of X for the discrete random variable X,

② $f_X(x; \theta_1, \theta_2, \ldots, \theta_k)$ is the probability density function of X for the continuous random variable,

③ $X_1, X_2, \ldots, X_n$ are the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively such that

(A) $X_1, X_2, \ldots, X_n$ are mutually independent random variables,

(B) the probability distribution function of each X_i is same as X $\forall\, i = 1(1)n$, i.e., for discrete random variable, the probability mass function of X_i is same as that of X whereas the probability density function of each X_i is same as that of X for the case of continuous random variable.

In maximum likelihood method, we want to find the unknown parameters $\theta_1, \theta_2, \ldots, \theta_K$ as a function of the sample values $x_1, x_2, \ldots, x_n$ in such a way that the parameters $\theta_1, \theta_2, \ldots, \theta_K$ maximize the likelihood function L, i.e., we want to find the maximum value of L with respect to the unknown parameter.

Therefore, if L is maximum at $\theta_i = \hat{\theta}_i, \forall i = 1(1)k$, then we get

$$\left.\frac{\partial L}{\partial \theta_i}\right|_{\Theta = \hat{\Theta}} = 0 \quad \forall i = 1(1)k,$$

$$\left. d^2 L \right|_{\Theta = \hat{\Theta}} < 0,$$

where $\Theta = (\theta_1, \theta_2, \ldots, \theta_K)$ and

$$\hat{\Theta} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_K)$$

Now it is simple to check that

$$\frac{\partial L}{\partial \theta_i} = 0 \implies \frac{\partial}{\partial \theta_i}(\log L) = 0$$

and consequently one may consider the function $\log L$ instead of L.

Application of maximum likelihood Method

(1) Find the maximum likelihood estimate of p for a Binomial (N, p) population.

Answer: Let $X = B(N, p)$ and consider a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of X. Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independend random variables and $X_i = B(N, p)$ for all $i = 1(1)n$. So, we can write

$$P(X_i = x_i) = \binom{N}{x_i} p^{x_i} (1-p)^{N-x_i} \quad \forall i = 1(1)n$$

Therefore the likelihood function for the sample $x_1, x_2, \ldots, x_n$ can be written as

$$L = P(X_1 = x_1)\, P(X_2 = x_2) \cdots P(X_n = x_n)$$

$$= \underbrace{\binom{N}{x_1} p^{x_1} (1-p)^{N-x_1}}\ \underbrace{\binom{N}{x_2} p^{x_2} (1-p)^{N-x_2}} \cdots$$

$$\underbrace{\binom{N}{x_n} p^{x_n} (1-p)^{N-x_n}}$$

$$L = \underbrace{\binom{N}{x_1}\binom{N}{x_2}\cdots\binom{N}{x_n}}_{A} \, p^{x_1+x_2+\cdots+x_n} (1-p)^{N-x_1+N-x_2+\cdots+N-x_n}$$

$$= A\, p^{x_1+x_2+\cdots+x_n} (1-p)^{nN-(x_1+x_2+\cdots+x_n)}$$

where $A = \binom{N}{x_1}\binom{N}{x_2}\cdots\binom{N}{x_n}$

$$= A\, p^{n\bar{x}} (1-p)^{nN-n\bar{x}} \quad \text{where } \bar{x} = \frac{1}{n}(x_1+x_2+\cdots+x_n)$$

$$= A\, p^{n\bar{x}} (1-p)^{n(N-\bar{x})}$$

$$\log L = \log A + n\bar{x}\log p + n(N-\bar{x})\log(1-p)$$

Here the unknown parameter is p only and consequently we want to maximize L or $\log L$ with respect to the unknown parameter p.

$$\frac{\partial(\log L)}{\partial p} = \underbrace{\frac{\partial}{\partial p}(\log A)}_{0} + \frac{\partial}{\partial p}(n\bar{x}\log p) + \frac{\partial}{\partial p}\{n(N-\bar{x})\log(1-p)\}$$

$$= \frac{n\bar{x}}{p} + \frac{n(N-\bar{x})}{(1-p)}(-1) = \frac{n\bar{x}}{p} - \frac{n(N-\bar{x})}{1-p}$$

$$\frac{\partial}{\partial p}(\log L)\Big|_{p=\hat{p}} = 0$$

$$\Downarrow$$

$$\frac{n\bar{x}}{\hat{p}} - \frac{n(N-\bar{x})}{1-\hat{p}} = 0$$

$$\Downarrow$$

$$\frac{\bar{x}}{\hat{p}} = \frac{N-\bar{x}}{1-\hat{p}}$$

$$\Downarrow$$

$$\bar{x}(1-\hat{p}) = \hat{p}(N-\bar{x})$$

$$\Downarrow$$

$$\bar{x} = \bar{x}\hat{p} + \hat{p}N - \hat{p}\bar{x} = \hat{p}N$$

$$\Downarrow$$

$$\hat{p} = \frac{\bar{x}}{N} = \frac{x_1 + x_2 + \cdots + x_n}{nN} \quad \text{———} \quad (1)$$

Therefore $\hat{p}$ given by equation (1) is the maximum likelihood estimate of p. To find the condition of the maximum/minimum/stationary one can use the formula

$$\checkmark \; d^2L = \sum_{i=1}^{k} \frac{\partial^2 L}{\partial \theta_i^2}(d\theta_i)^2 + \sum_{\substack{i=1 \\ i \neq j}}^{k} \sum_{j=1}^{k} \frac{\partial^2 L}{\partial \theta_i \partial \theta_j} d\theta_i \, d\theta_j$$

(2) Poisson - μ population : Find the maximum likelihood estimate of μ for a Poisson μ population.

Answer: Let $X = \mathcal{P}(\mu)$ and consider a sample of size n: $x_1, x_2, \dots, x_n$ drawn from the population of X. Let $X_1, X_2, \dots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \dots, x_n$ respectively. Therefore, $X_1, X_2, \dots, X_n$ are mutually independent random variables and $X_i = \mathcal{P}(\mu)$ $\forall\ i = 1(1)n$. So,

$$P(X_i = x_i) = \frac{e^{-\mu}\mu^{x_i}}{(x_i)!} \quad \forall\ x_i = 0, 1, 2, \dots$$

Therefore the likelihood function of the sample $x_1, x_2, \dots, x_n$ can be written as

$$L = P(X_1 = x_1)\, P(X_2 = x_2) \cdots P(X_n = x_n)$$

$$= \frac{e^{-\mu}\mu^{x_1}}{(x_1)!} \cdot \frac{e^{-\mu}\mu^{x_2}}{(x_2)!} \cdots \frac{e^{-\mu}\mu^{x_n}}{(x_n)!}$$

$$= \frac{e^{-n\mu}\mu^{x_1 + x_2 + \dots + x_n}}{(x_1)!\,(x_2)! \cdots (x_n)!} = \frac{e^{-n\mu}\mu^{n\bar{x}}}{(x_1)!\,(x_2)! \cdots (x_n)!}$$

$$L = A e^{-n\mu} \mu^{n\bar{x}} \qquad \text{where } \bar{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

$$A = \frac{1}{(x_1)!\,(x_2)!\cdots(x_n)!}$$

$$\Rightarrow \log L = \log A - n\mu + n\bar{x}\log\mu$$

$$\frac{\partial(\log L)}{\partial\mu} = -n + \frac{n\bar{x}}{\mu}$$

$$\frac{\partial}{\partial\mu}(\log L)\Big|_{\mu=\hat{\mu}} = 0 \Rightarrow -n + \frac{n\bar{x}}{\hat{\mu}} = 0$$

$$\Rightarrow \hat{\mu} = \bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} \qquad \text{———} \; (1)$$

$\hat{\mu}$ as given in equation (1) is the maximum likelihood estimate of μ on the basis of the sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the Poisson μ populatation. From equation (1) we see that $\hat{\mu}$ is a consistent and unbiased estimate of μ.

(3) <u>Normal Population – $N(m,\sigma)$</u>: Let $X = N(m,\sigma)$ and consider a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of X. Let $X_1, X_2, \ldots, X_n$ be the random variables crrosponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent and the probability density function of each X_i is same as X for all $i = 1(1)n$. Consequently we have

$X_i = N(m,\sigma)$ $\forall\, i = 1(1)n$, i.e.,

$$f_{X_i}(x_i; m,\sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x_i - m}{\sigma}\right)^2}, \quad -\infty < x_i < \infty,\ i = 1(1)n$$

Therefore the likelihood function for the sample values $x_1, x_2, \ldots, x_n$ can be written as

$$L = f_{X_1}(x_1; m,\sigma)\, f_{X_2}(x_2; m,\sigma) \cdots f_{X_n}(x_n; m,\sigma)$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x_1 - m}{\sigma}\right)^2} \cdot \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x_2 - m}{\sigma}\right)^2} \cdots$$

$$\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x_n - m}{\sigma}\right)^2}$$

$$= \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\frac{1}{2}\left[\left(\frac{x_1 - m}{\sigma}\right)^2 + \left(\frac{x_2 - m}{\sigma}\right)^2 + \cdots + \left(\frac{x_n - m}{\sigma}\right)^2\right]}$$

$$L = (2\pi)^{-\frac{n}{2}} \sigma^{-n} e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n} (x_i - m)^2\right]}$$

$$\Rightarrow \log L = -\frac{n}{2}\ln(2\pi) - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2$$

Let L attains its maximum value at $m = \hat{m}$ and $\sigma = \hat{\sigma}$

$$\therefore \frac{\partial}{\partial m}(\log L)\Big|_{m=\hat{m},\, \sigma=\hat{\sigma}} = 0 \quad \text{———} \quad (1)$$

$$\frac{\partial}{\partial \sigma}(\log L)\Big|_{m=\hat{m},\, \sigma=\hat{\sigma}} = 0 \quad \text{———} \quad (2)$$

Now $\displaystyle \frac{\partial}{\partial m}(\log L) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i - m)(-1)$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - m)$$

$$= \frac{1}{\sigma^2}\left[\sum_{i=1}^{n} x_i - mn\right]$$

$$= \frac{1}{\sigma^2}\left[n\bar{x} - mn\right] = \frac{n}{\sigma^2}\left[\bar{x} - m\right] \text{ — } (3)$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\frac{\partial}{\partial \sigma}(\log L) = -\frac{n}{\sigma} - \frac{1}{2}\cdot(-2)\,\sigma^{-2-1}\sum_{i=1}^{n}(x_i - m)^2$$

$$= -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - m)^2 \quad \text{———} \quad (4)$$

From 1, 2, 3 & 4 we get

$$\frac{n}{\hat{\sigma}^2}[\bar{x} - \hat{m}] = 0 \qquad (5)$$

$$-\frac{n}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^3}\sum_{i=1}^{n}(x_i - \hat{m})^2 = 0 \qquad (6)$$

From 5 we get $\boxed{\hat{m} = \bar{x}}$ (7)

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{m})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = s^2$$

Therefore, likelihood estimate of m is $\bar{x}$, the sample mean which is a consistent and unbiased estimate of the population of mean and the likelihood estimate of σ^2 is s^2, the sample variance which is a consistent estimate of the population variance but it is not an unbiased estimate of population variance. Specifically, sample variance is negatively biased estimate of the population variance.

Problem: If the value of m is known, find the maximum likelihood estimate of σ^2 for a Normal (m, σ) population and show that the estimate is unbiased and consistent.

Answer: Let $X = N(m, \sigma)$ and consider a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of the random variable X. Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the samble values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent and probability density function of each X_i is same as X for all $i = 1(1)n$. Consequently, we have

$X_i = N(m, \sigma)$ $\forall\, i = 1(1)n$, m is known, i.e.,

$$f_{X_i}(x_i; \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x_i - m)^2}, \quad -\infty < x_i < \infty$$

for $i = 1, 2, \ldots, n$

Therefore, the likelihood function L for the sample values $x_1, x_2, \dots, x_n$ can be written as

$$L = f_{X_1}(x_1;\sigma)\, f_{X_2}(x_2;\sigma) \cdots f_{X_n}(x_n;\sigma)$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x_1-m}{\sigma}\right)^2} \cdot \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x_2-m}{\sigma}\right)^2} \cdots$$

$$\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x_n-m}{\sigma}\right)^2}$$

$$= (2\pi)^{-\frac{n}{2}} \sigma^{-n} e^{-\frac{1}{2}\left[\left(\frac{x_1-m}{\sigma}\right)^2 + \left(\frac{x_2-m}{\sigma}\right)^2 + \cdots + \left(\frac{x_n-m}{\sigma}\right)^2\right]}$$

$$= (2\pi)^{-\frac{n}{2}} \sigma^{-n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2}$$

$$\log L = -\frac{n}{2}\ln 2\pi - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2$$

Let $\hat{\sigma}$ be the maximum likelihood estimate of σ.

Therefore, we have

$$\frac{\partial}{\partial\sigma}(\ln L)\Big|_{\sigma=\hat{\sigma}} = 0 \quad \text{——— (1)}$$

$$\frac{\partial}{\partial\sigma}(\ln L) = -\frac{n}{\sigma} - \frac{1}{2}(-2)\sigma^{-3}\sum_{i=1}^{n}(x_i-m)^2$$

$$= -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-m)^2 \quad \text{——— (2)}$$

From (1) & (2) we get

$$-\frac{n}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^3}\sum_{i=1}^{n}(x_i-m)^2 = 0$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 \quad \text{———} \quad (3)$$

The likelihood estimate $\hat{\sigma}$ of σ is given by equation (3) when the value of $\hat{\sigma}$ is given by equation (3). Now the random variable A corresponding to the statistic $\hat{\sigma}^2$ is given by

$$A = \frac{1}{n}\sum (X_i - m)^2$$

Now as $X_i = N(m, \sigma)$, $E(X_i) = m$ and

$var(X_i) = \sigma^2 \quad \forall\, i = 1(1)n$

$$\therefore\ A = \frac{1}{n}\sum_{i=1}^{n}[X_i - E(X_i)]^2$$

$$\Rightarrow E(A) = \frac{1}{n}\sum_{i=1}^{n} E(X_i - E(X_i))^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n}\cdot n\sigma^2$$

$\Rightarrow E(A) = \sigma^2 \Rightarrow \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2$ is an unbiased estimate of $\sigma^2 \Rightarrow \hat{\sigma}^2$ is an unbiased estimate of $\sigma^2 \Rightarrow$ Maximum likelihood estimate of σ^2 is an unbiased estimate of σ^2.

$$A = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2m x_i + m^2)$$

$$= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2m\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) + \frac{1}{n} \times n m^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2m \frac{1}{n}\sum_{i=1}^{n} x_i + m^2$$

$$= A_2 - 2m A_1 + m^2$$

We know that A_k is the consistent estimate of α_k

i.e., $A_k \xrightarrow[mp]{} \alpha_k$ as $n \to \infty$

$\therefore A_2 \xrightarrow[mp]{} \alpha_2$, $A_1 \xrightarrow[mp]{} \alpha_1$, $m^2 \xrightarrow[mp]{} m^2$ as $n \to \infty$

$$\Rightarrow A_2 - 2m A_1 + m^2 \xrightarrow[mp]{} \alpha_2 - 2m \alpha_1 + m^2 \text{ as } n \to \infty \quad (4)$$

Now $\alpha_2 - 2m\alpha_1 + m^2 = \alpha_2 - 2m m + m^2$ [as $\alpha_1 = m$]

$$= \alpha_2 - m^2$$

$$= \alpha_2 - (\alpha_1)^2 = \mu_2 = \sigma^2$$

$\therefore A_2 - 2mA_1 + m^2 \xrightarrow[mp]{} \sigma^2$ as $n \to \infty$

$\Rightarrow A \xrightarrow[mp]{} \sigma^2$ as $n \to \infty$

$\Rightarrow \hat{\sigma}^2$ is the consistent estimate of σ^2

Problem: Use the maximum likelihood method to estimate the parameter α of a continuous population having density function $(1+\alpha)x^{\alpha}$ for $0 < x < 1$.

Answer: Consider a sample $x_1, x_2, \ldots, x_n$ of size n drawn from the population of X, where the probability density function of X is given by

$$f_X(x;\alpha) = (1+\alpha)\,x^{\alpha} \quad \text{for } 0 < x < 1$$
$$= 0 \quad \text{Otherwise}$$

Here we assume that the following condition holds good: $0 < x_i < 1 \quad \forall\, i = 1(1)n$.

Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent random variable and the probability density function of X_i is same as X for all $i = 1(1)n$.

Therefore, the likelihood function L for the sample values $x_1, x_2, \ldots, x_n$ can be written as

$$L = f_{X_1}(x_1; \alpha)\, f_{X_2}(x_2; \alpha) \cdots f_{X_n}(x_n; \alpha)$$

$$= (1+\alpha)\, x_1^{\alpha}\, (1+\alpha)\, x_2^{\alpha} \cdots (1+\alpha)\, x_n^{\alpha}$$

$$= (1+\alpha)^n (x_1 x_2 \cdots x_n)^{\alpha}$$

$$\ln L = n \ln(1+\alpha) + \alpha \ln(x_1 x_2 \cdots x_n)$$

$$\frac{\partial}{\partial \alpha}(\ln L) = \frac{n}{1+\alpha} + \ln(x_1 x_2 \cdots x_n)$$

Let $\hat{\alpha}$ be the maximum likelihood estimate of α.

$$\therefore \left.\frac{\partial}{\partial \alpha}(\ln L)\right|_{\hat{\alpha}} = 0$$

$$\Rightarrow \frac{n}{1+\hat{\alpha}} + \ln(x_1 x_2 \cdots x_n) = 0$$

$$\Rightarrow \frac{n}{1+\hat{\alpha}} = -\ln(x_1 x_2 \cdots x_n)$$

$$\Rightarrow 1+\hat{\alpha} = -\frac{n}{\ln(x_1 x_2 \cdots x_n)} = \frac{n}{\ln(x_1 x_2 \cdots x_n)^{-1}}$$

$$\Rightarrow \hat{\alpha} = \frac{n}{\ln(x_1 x_2 \cdots x_n)^{-1}} - 1$$

Problem: Use the maximum likelihood method to prove that the estimate of the parameter α of a population having density function $2(\alpha - x)/\alpha^2$ $(0 < x < \alpha)$ for a sample x of unit size is $2x$. Show that the estimate is biased.

Answer: Here the likelihood function L for a sample x of unit size is given by

$$L = \frac{2(\alpha - x)}{\alpha^2}$$

$$\Rightarrow \ln L = \ln 2 + \ln(\alpha - x) - 2\ln\alpha$$

$$\Rightarrow \frac{\partial}{\partial\alpha}(\ln L) = \frac{1}{\alpha - x} - \frac{2}{\alpha} \quad \text{—— (1)}$$

Let $\hat{\alpha}$ be the maximum likelihood estimate of α for the sample x of unit size.

$$\left.\frac{\partial}{\partial\alpha}(\ln L)\right|_{\alpha = \hat{\alpha}} = 0$$

$$\Rightarrow \frac{1}{\hat{\alpha} - x} - \frac{2}{\hat{\alpha}} = 0 \Rightarrow \frac{1}{\hat{\alpha} - x} = \frac{2}{\hat{\alpha}} \Rightarrow \hat{\alpha} = 2(\hat{\alpha} - x)$$

$$\Rightarrow \hat{\alpha} = 2\hat{\alpha} - 2x \Rightarrow 2\hat{\alpha} - \hat{\alpha} = 2x \Rightarrow \boxed{\hat{\alpha} = 2x}$$

The random variable corresponding to $\hat{\alpha}$ is given by

$A = 2X$

$$\Rightarrow E(A) = E(2X) = 2E(X) = 2\int_0^{\alpha} \frac{2(\alpha - x)}{\alpha^2}\, x\, dx$$

$$= \frac{4}{\alpha^2}\int_0^{\alpha} (\alpha x - x^2)\, dx = \frac{4}{\alpha^2}\left[\frac{\alpha\,\alpha^2}{2} - \frac{\alpha^3}{3}\right]$$

$$= \frac{4}{\alpha^2}\left[\frac{\alpha^3}{2} - \frac{\alpha^3}{3}\right] = \frac{4}{\alpha^2}\cdot\frac{\alpha^3}{6} = \frac{2}{3}\alpha \neq \alpha$$

$\Rightarrow E(A) \neq \alpha$

$\Rightarrow \hat{\alpha}$ is the biased estimate of α and it is negatively biased because $E(A) - \alpha = \frac{2}{3}\alpha - \alpha = -\frac{\alpha}{3} < 0$.

<u>Problem</u>: Use the maximum likelihood method to estimate the parameter p of a discrete population having probability mass function

$P(X = i) = p(1-p)^{i-1}$, $i = 1, 2, 3, \ldots$, $0 < p < 1$.

Use a sample of size n as $x_1, x_2, \ldots, x_n$.

<u>Answer</u>: Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually

independent random variables and probability mass function of each X_i is same as X, i.e.,
$P(X_i = x_i) = p\,q^{x_i - 1}$, $q = 1-p$, $i = 1(1)n$

Therefore, the likelihood function L is given by

$$L = P(X_1 = x_1)\,P(X_2 = x_2)\cdots P(X_n = x_n)$$

$$= p\,q^{x_1-1}\cdot p\,q^{x_2-1}\cdots p\,q^{x_n-1}$$

$$= p^n\,q^{x_1-1+x_2-1+\cdots+x_n-1}$$

$$= p^n\,q^{x_1+x_2+\cdots+x_n-n}$$

$$= p^n\,q^{n\bar{x}-n} \quad , \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$= p^n(1-p)^{n\bar{x}-n}$$

$$\log L = n\ln p + (n\bar{x}-n)\ln(1-p)$$

$$\frac{\partial}{\partial p}(\log L) = \frac{n}{p} + \frac{n\bar{x}-n}{1-p}\cdot(-1) = \frac{n}{p} - \frac{n\bar{x}-n}{1-p}$$

Let $\hat{p}$ is the maximum likelihood estimate of p.

$$\therefore \left.\frac{\partial}{\partial p}(\log L)\right|_{p=\hat{p}} = 0 \Rightarrow \frac{n}{\hat{p}} - \frac{n\bar{x}-n}{1-\hat{p}} = 0$$

$$\Rightarrow \frac{1}{\hat{p}} = \frac{\bar{x}-1}{1-\hat{p}} \Rightarrow 1-\hat{p} = \hat{p}(\bar{x}-1) \Rightarrow 1 = \hat{p}(\bar{x}-1) + \hat{p}$$

$$\Rightarrow 1 = \hat{p}(\bar{x}-1+1) = \hat{p}\,\bar{x} \Rightarrow \boxed{\hat{p} = \frac{1}{\bar{x}}}$$

$\therefore$ the maximum likelihood estimate of p for the sample $x_1, x_2, \ldots, x_n$ is

$$\frac{n}{x_1 + x_2 + \cdots + x_n}$$

<u>Problem</u>: Use the maximum likelihood method to estimate the parameter μ of a discrete population having probability mass function

$$P(X=i) = \frac{1}{1+\mu}\left(\frac{\mu}{1+\mu}\right)^{i}, \quad i = 0, 1, 2, \ldots, \quad \mu > 0.$$

Use a sample of size n as $x_1, x_2, \ldots, x_n$.

<u>Answer</u>: Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent random variables and each X_i has the same probability mass function as X. Therefor the likelihood function L for the given sample is

$$L = P(X_1 = x_1)\, P(X_2 = x_2) \cdots P(X_n = x_n)$$

$$= \frac{1}{1+\mu}\left(\frac{\mu}{1+\mu}\right)^{x_1} \frac{1}{1+\mu}\left(\frac{\mu}{1+\mu}\right)^{x_2} \cdots \frac{1}{1+\mu}\left(\frac{\mu}{1+\mu}\right)^{x_n}$$

$$= \left(\frac{1}{1+\mu}\right)^{n}\left(\frac{\mu}{1+\mu}\right)^{x_1 + x_2 + \cdots + x_n}$$

$$= \left(\frac{1}{1+\mu}\right)^{n}\left(\frac{\mu}{1+\mu}\right)^{n\bar{x}} \qquad \text{where } \bar{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

$$= \frac{\mu^{n\bar{x}}}{(1+\mu)^{n\bar{x}+n}}$$

$$L = \frac{\mu^{n\bar{x}}}{(1+\mu)^{n\bar{x}+n}}$$

$$\Rightarrow \ln L = n\bar{x} \ln \mu - (n\bar{x}+n) \ln(1+\mu)$$

$$\Rightarrow \frac{\partial}{\partial \mu}(\ln L) = \frac{n\bar{x}}{\mu} - \frac{n\bar{x}+n}{1+\mu}$$

Let $\hat{\mu}$ be the maximum likelihood estimate of μ

$$\therefore \left.\frac{\partial}{\partial \mu}(\ln L)\right|_{\mu=\hat{\mu}} = 0$$

$$\Rightarrow \frac{n\bar{x}}{\hat{\mu}} - \frac{n\bar{x}+n}{1+\hat{\mu}} = 0$$

$$\Rightarrow \frac{n\bar{x}}{\hat{\mu}} = \frac{n(\bar{x}+1)}{1+\hat{\mu}}$$

$$\Rightarrow \frac{\bar{x}}{\hat{\mu}} = \frac{\bar{x}+1}{1+\hat{\mu}}$$

$$\Rightarrow \bar{x}(1+\hat{\mu}) = \hat{\mu}(\bar{x}+1)$$

$$\Rightarrow \bar{x} = \hat{\mu} \Rightarrow \boxed{\hat{\mu} = \bar{x}}$$

Therefore the maximum likelihood estimate of μ

= the sample mean which is a consistent

and unbiased estimate of population mean.

Problem: A population is defined by the following probability density function

$$f_X(x; a) = \frac{x^{l-1} e^{-x/a}}{\Gamma(l)\, a^l}, \quad 0 < x < \infty,$$

where l is a known constant. Use the maximum likelihood method to estimate the unknown constant a on the basis of a sample of size n: $x_1, x_2, \ldots, x_n$. Show that the estimate is consistent and unbiased.

Answer: Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent random variables and the probability distribution of each X_i is same as that of the population. Therefore the likelihood function L for the given sample can be written as

$$L = f_{X_1}(x_1; a)\, f_{X_2}(x_2; a) \cdots f_{X_n}(x_n; a)$$

$$= \frac{x_1^{l-1} e^{-x_1/a}}{\Gamma(l)\, a^l} \cdot \frac{x_2^{l-1} e^{-x_2/a}}{\Gamma(l)\, a^l} \cdots \frac{x_n^{l-1} e^{-x_n/a}}{\Gamma(l)\, a^l}$$

$$L = \frac{(x_1 x_2 \cdots x_n)^{l-1} e^{-\frac{1}{a}(x_1+x_2+\cdots+x_n)}}{(\Gamma(l))^n a^{nl}}$$

$$= \frac{(x_1 x_2 \cdots x_n)^{l-1}}{(\Gamma(l))^n} \cdot \frac{e^{-\frac{n\bar{x}}{a}}}{a^{nl}}, \quad \bar{x} = \frac{1}{n}(x_1+x_2+\cdots+x_n)$$

$$= B e^{-\frac{n\bar{x}}{a}} a^{-nl}, \quad B = \frac{(x_1 x_2 \cdots x_n)^{l-1}}{(\Gamma(l))^n} \rightarrow \text{independent of } a$$

$$\Rightarrow \ln L = \ln B - \frac{n\bar{x}}{a} - nl \ln a$$

$$\Rightarrow \frac{\partial}{\partial a}(\ln L) = \frac{n\bar{x}}{a^2} - \frac{nl}{a}$$

Let $\hat{a}$ be the maximum likelihood estimate of a

$$\therefore \left.\frac{\partial}{\partial a}(\ln L)\right|_{a=\hat{a}} = 0$$

$$\Rightarrow \frac{n\bar{x}}{\hat{a}^2} - \frac{nl}{\hat{a}} = 0$$

$$\Rightarrow \frac{\bar{x}}{\hat{a}} - l = 0$$

$$\Rightarrow \hat{a} = \frac{\bar{x}}{l}$$

We know that $\bar{x}$ is a consistent and unbiased estimate of the population mean $E(x)$

$\Rightarrow \frac{\bar{x}}{l}$ is a consistent and unbiased estimate of $\frac{E(x)}{l}$

$\Rightarrow \hat{a}$ is a consitent and unbiased estimate of $\frac{E(x)}{l} = \frac{al}{l} = a$

[It is simple to check that $E(X) = al$]

Interval Estimation

Let α be a population parameter and ε be a real number such that $0 < \varepsilon < 1$. consider a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of the random variable X.

Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Therefore, $X_1, X_2, \ldots, X_n$ are mutually independent random variables and the probability distribution of each X_i is same as X. If there exist two statistics

$$a = a(x_1, x_2, \ldots, x_n) \text{ and } b = b(x_1, x_2, \ldots, x_n)$$

such that

$$P(A < \alpha < B) = 1 - \varepsilon$$

then the interval (a, b) is called the confidence interval of α with confidence coefficient $1-\varepsilon$, where A and B are the random variable corresponding to the statistics a and b respectively, i.e.,

$$A = a(X_1, X_2, \ldots, X_n) \text{ and } B = b(X_1, X_2, \ldots, X_n)$$

Problem-1 : find the confidence interval of m for a normal (m, σ) population when σ is known.

Answer: Consider a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of $X = N(m, \sigma)$, the value of σ is given. Let $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively. Then the random variables $X_1, X_2, \ldots, X_n$ are mutually independent and $X_i = N(m, \sigma)$ $\forall\, i = 1(1)n$. Then

$\bar{X} = N(m, \frac{\sigma}{\sqrt{n}})$ where $\bar{X} = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$

Here the appropriate statistic is

$$u = \frac{\bar{x} - m}{\sigma/\sqrt{n}}$$

whose sampling distribution is $N(0,1)$

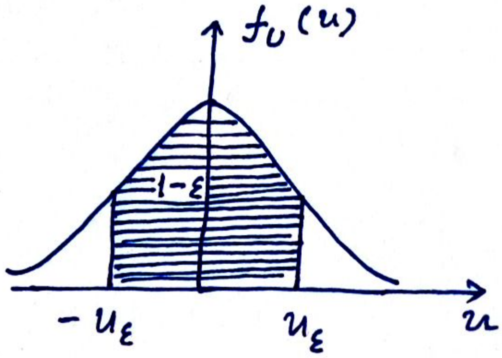
i.e., $U = \dfrac{\bar{X} - m}{\sigma/\sqrt{n}} = N(0,1)$

$\therefore \quad U = N(0, 1)$

For given ε, $0<\varepsilon<1$, take two real numbers $\pm u_\varepsilon$ such that

$$P(-u_\varepsilon < U < u_\varepsilon) = 1-\varepsilon \qquad \text{——— (1)}$$

$$\Rightarrow \int_{-u_\varepsilon}^{u_\varepsilon} f_U(u)\,du = 1-\varepsilon$$



$$\Rightarrow 2\int_{0}^{u_\varepsilon} f_U(u)\,du = 1-\varepsilon$$

$$\Rightarrow \int_{0}^{u_\varepsilon} f_U(u)\,du = \frac{1-\varepsilon}{2}$$

$$\Rightarrow P(0<U<u_\varepsilon) = \frac{1-\varepsilon}{2}$$

$$\Rightarrow P(0<U<\infty) - P(U>u_\varepsilon) = \frac{1-\varepsilon}{2}$$

$$\Rightarrow \frac{1}{2} - P(U>u_\varepsilon) = \frac{1-\varepsilon}{2} \qquad \left[\because P(-\infty<U<\infty)=1 \right.$$

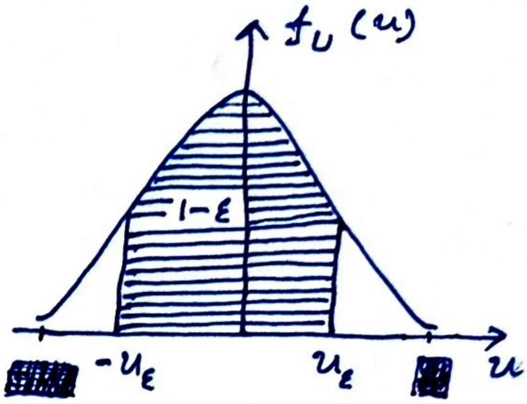
$$\Rightarrow 2P(0<U<\infty) = 1$$

$$\Rightarrow P(U>u_\varepsilon) = \frac{1}{2} - \frac{1-\varepsilon}{2} = \frac{\varepsilon}{2} \qquad \left. \Rightarrow P(0<U<\infty) = \frac{1}{2}\right]$$

$$\Rightarrow \boxed{P(U>u_\varepsilon) = \frac{\varepsilon}{2}} \qquad \text{——— (2)}$$

$$\Rightarrow \int_{u_\varepsilon}^{\infty} f_U(u)\,du = \frac{\varepsilon}{2} \rightarrow$$ this is an equation for the unknowns u_ε and this equation determines u_ε, i.e., for given ε, equation (2) determines the constant u_ε.

From ① we get



$$P(-u_\varepsilon < U < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|U| < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P\left(\left|\frac{\bar{x}-m}{\sigma/\sqrt{n}}\right| < u_\varepsilon\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\left|\frac{\bar{x}-m}{(\sigma/\sqrt{n})}\right| < u_\varepsilon\right) = 1-\varepsilon$$

$$\Rightarrow P\left(|\bar{x}-m| < \frac{\sigma u_\varepsilon}{\sqrt{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(|m-\bar{x}| < \frac{\sigma u_\varepsilon}{\sqrt{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(-\frac{\sigma u_\varepsilon}{\sqrt{n}} < m-\bar{x} < \frac{\sigma u_\varepsilon}{\sqrt{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\bar{x} - \frac{\sigma u_\varepsilon}{\sqrt{n}} < m < \bar{x} + \frac{\sigma u_\varepsilon}{\sqrt{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P(A < m < B) = 1-\varepsilon$$

where $A = \bar{x} - \dfrac{\sigma u_\varepsilon}{\sqrt{n}}$, $B = \bar{x} + \dfrac{\sigma u_\varepsilon}{\sqrt{n}}$

and u_ε is given by equation ①

$\therefore$ (a, b) is the confidence interval of m with confidence coefficient $1-\varepsilon$, where

$a = \bar{x} - \dfrac{\sigma u_\varepsilon}{\sqrt{n}}$ and $b = \bar{x} + \dfrac{\sigma u_\varepsilon}{\sqrt{n}}$.

Problem - 2 : Find the confidence interval of m for a normal (m, σ) population when σ is unknown.

Answer : Here the appropriate statistic is

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

whose sampling distribution is t - distribution with $(n-1)$ degrees of freedom, where

① $\bar{x} = \dfrac{x_1 + x_2 + \cdots + x_n}{n}$

② $s^2 = \dfrac{1}{n-1}\left\{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2\right\}$

③ $x_1, x_2, \ldots, x_n$ be a sample of size n drawn from $N(m, \sigma)$ population

If T be the random corresponding to the statistic t, then

$$T = \frac{\bar{X} - m}{(S_1/\sqrt{n})}$$

where ① $\bar{X} = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$

② $S'^2 = \dfrac{1}{n-1}\left\{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2\right\}$

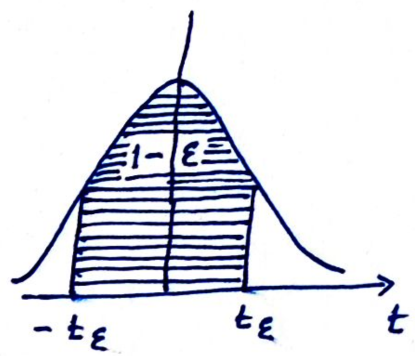
③ $X_1, X_2, \ldots, X_n$ are the random variables corresponding to the sample values

$x_1, x_2, \ldots, x_n$ respectively

(4) $X_1, X_2, \ldots, X_n$ are mutually independent random variables and $X_i = N(m, \sigma)$

For given ε, $0<\varepsilon<1$, we choose two real numbers $-t_\varepsilon$ and t_ε ($t_\varepsilon > 0$) such that

$$P(-t_\varepsilon < T < t_\varepsilon) = 1-\varepsilon \quad \text{——— (1)}$$



$$\Rightarrow P(|T| < t_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow 1 - P(|T| \geq t_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|T| \geq t_\varepsilon) = \varepsilon$$

$$\Rightarrow P(|T| > t_\varepsilon) = \varepsilon$$

$$\Rightarrow P(T > t_\varepsilon) + P(T < -t_\varepsilon) = \varepsilon$$

$$\Rightarrow P(T > t_\varepsilon) + P(T > t_\varepsilon) = \varepsilon$$ [$\because$ The distribution function of T is symmetrical about the origin]

$$\Rightarrow 2P(T > t_\varepsilon) = \varepsilon$$

$$\Rightarrow P(T > t_\varepsilon) = \frac{\varepsilon}{2} \quad \text{——— (2)}$$

$$\Rightarrow \int_{t_\varepsilon}^{\infty} f_T(t)\,dt = \frac{\varepsilon}{2}$$ → this is an equation for the unknown t_ε and this equation determines t_ε, i.e., for given ε, $0<\varepsilon<1$, the equation (2) determines the constant t_ε.

From ① we get

$$P(-t_\varepsilon < T < t_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|T| < t_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P\left(\left|\frac{\bar{x}-m}{s_1/\sqrt{n}}\right| < t_\varepsilon\right) = 1-\varepsilon$$

$$\Rightarrow P\left(|\bar{x}-m| < \frac{t_\varepsilon s_1}{\sqrt{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(|m-\bar{x}| < \frac{t_\varepsilon s_1}{\sqrt{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(-\frac{t_\varepsilon s_1}{\sqrt{n}} < m-\bar{x} < \frac{t_\varepsilon s_1}{\sqrt{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\bar{x}-\frac{t_\varepsilon s_1}{\sqrt{n}} < m < \bar{x}+\frac{t_\varepsilon s_1}{\sqrt{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P(A < m < B) = 1-\varepsilon$$

where

$$A = \bar{X} - \frac{t_\varepsilon s_1}{\sqrt{n}} \quad \text{and} \quad B = \bar{X} + \frac{t_\varepsilon s_1}{\sqrt{n}}$$

∴ The confidence interval of m at the confidence level $1-\varepsilon$ is (a, b), where

$$a = \bar{x} - \frac{t_\varepsilon s}{\sqrt{n}} \quad \text{and} \quad b = \bar{x} + \frac{t_\varepsilon s}{\sqrt{n}}$$

and t_ε can be determine through equation ②.

<u>Problem-3</u> : Find the confidence interval of σ for a normal (m, σ) population.

<u>Answer</u> : Here the appropriate statistic is

$$\chi^2 = \frac{n s^2}{\sigma^2},$$

which is χ^2- distributed with $(n-1)$ degrees freedom,

where

① $s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$

② $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

③ $x_1, x_2, \ldots, x_n$ be a sample of size n drawn from the population of $X = N(m, \sigma)$

The random variable corresponding to the statistic χ^2 is given by

$$T = \frac{n S_1^2}{\sigma^2},$$

where

① $S_1^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$

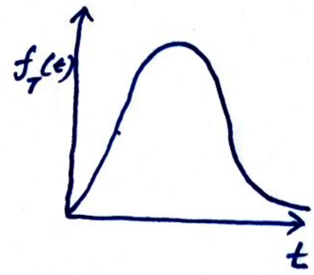
② $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$

③ $X_1, X_2, \ldots, X_n$ are the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively

④ $x_1, x_2, \ldots, x_n$ are mutually independent and

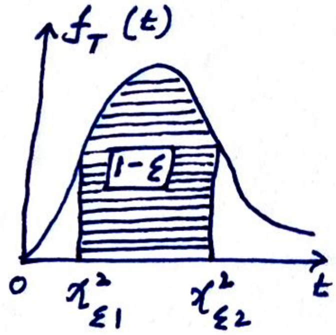
⑤ $x_i = N(m, \sigma) \quad \forall\, i = 1(1)n$

Therefore T is χ^2 distributed with $(n-1)$ degrees of freedom.



Choose two strictly positive real numbers $\chi^2_{\varepsilon 1}$ and $\chi^2_{\varepsilon 2}$ such that

$$P(\chi^2_{\varepsilon 1} < T < \chi^2_{\varepsilon 2}) = 1 - \varepsilon \qquad \text{①}$$



$$\Rightarrow P\left(\chi^2_{\varepsilon 1} < \frac{n s_1^2}{\sigma^2} < \chi^2_{\varepsilon 2}\right) = 1 - \varepsilon$$

$$\Rightarrow P\left(\frac{1}{\chi^2_{\varepsilon 1}} > \frac{\sigma^2}{n s_1^2} > \frac{1}{\chi^2_{\varepsilon 2}}\right) = 1 - \varepsilon$$

$$\Rightarrow P\left(\frac{1}{\chi^2_{\varepsilon 2}} < \frac{\sigma^2}{n s_1^2} < \frac{1}{\chi^2_{\varepsilon 1}}\right) = 1 - \varepsilon$$

$$\Rightarrow P\left(\frac{n s_1^2}{\chi^2_{\varepsilon 2}} < \sigma^2 < \frac{n s_1^2}{\chi^2_{\varepsilon 1}}\right) = 1 - \varepsilon \qquad \text{(1A)}$$

$$\Rightarrow 1 - \left[P\left(\frac{n s_1^2}{\chi^2_{\varepsilon 2}} > \sigma^2\right) + P\left(\sigma^2 > \frac{n s_1^2}{\chi^2_{\varepsilon 1}}\right)\right] = 1 - \varepsilon$$

$$\Rightarrow P\left(\frac{n s_1^2}{\chi^2_{\varepsilon 2}} > \sigma^2\right) + P\left(\sigma^2 > \frac{n s_1^2}{\chi^2_{\varepsilon 1}}\right) = \varepsilon \qquad \text{②}$$

$$\downarrow \qquad\qquad\qquad \downarrow$$

$$\frac{n s_1^2}{\sigma^2} > \chi^2_{\varepsilon 2} \qquad\qquad \frac{n s_1^2}{\sigma^2} < \chi^2_{\varepsilon 1}$$

$$\downarrow \qquad\qquad\qquad \downarrow$$

$$\chi^2_{\varepsilon 2} < \frac{n s_1^2}{\sigma^2} < \infty \qquad\qquad 0 < \frac{n s_1^2}{\sigma^2} < \chi^2_{\varepsilon 1}$$

$$\Rightarrow P\left(\chi^2_{\varepsilon 2} < \frac{n s_1^2}{\sigma^2} < \infty\right) + P\left(0 < \frac{n s_1^2}{\sigma^2} < \chi^2_{\varepsilon 1}\right) = \varepsilon$$

$$\Rightarrow P(T > \chi^2_{\varepsilon 2}) + P(0 < T < \chi^2_{\varepsilon 1}) = \varepsilon \quad \text{——— ③}$$

It is not possible to get two unknowns from only one equation ③. So, for this case, the general prescription is to choose $\chi^2_{\varepsilon 1}$ and $\chi^2_{\varepsilon 2}$ from the equations

$$P(0 < T < \chi^2_{\varepsilon 1}) = \varepsilon/2 \quad \text{——— ④}$$

$$P(T > \chi^2_{\varepsilon 2}) = \varepsilon/2 \quad \text{——— ⑤}$$

From (1A) we get

$$P\left(\frac{n s_1^2}{\chi^2_{\varepsilon 2}} < \sigma^2 < \frac{n s_1^2}{\chi^2_{\varepsilon 1}}\right) = 1 - \varepsilon$$

$$\Rightarrow P\left(\sqrt{\frac{n s_1^2}{\chi^2_{\varepsilon 2}}} < \sigma < \sqrt{\frac{n s_1^2}{\chi^2_{\varepsilon 1}}}\right) = 1 - \varepsilon$$

$$\Rightarrow P(A < \sigma < B) = 1 - \varepsilon$$

where

$$A = \sqrt{\frac{n s_1^2}{\chi^2_{\varepsilon 2}}}, \quad B = \sqrt{\frac{n s_1^2}{\chi^2_{\varepsilon 1}}}$$

$\chi^2_{\varepsilon 1}$, $\chi^2_{\varepsilon 2}$ are given by the equations ④ and ⑤.

Therefore, the confidence interval of σ at the confidence level $1-\varepsilon$ is (a, b), where

$$a = \sqrt{\frac{n s^2}{\chi^2_{\varepsilon 2}}}, \quad b = \sqrt{\frac{n s^2}{\chi^2_{\varepsilon 1}}}$$

$\chi^2_{\varepsilon 1}$ and $\chi^2_{\varepsilon 2}$ are given by ④ and ⑤.

Problem-4: Find the approximate confidence interval of p for Binomial (n, p) population.

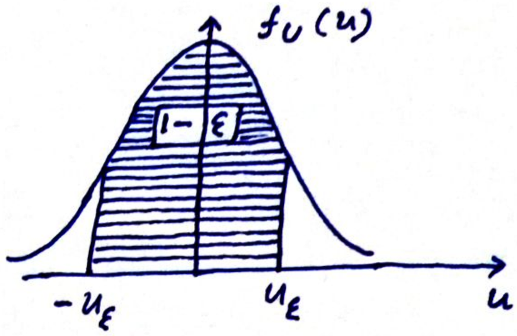
Answer-4: Let $X = B(n, p)$ and x be an observed value of X. For large sample, we know that

$\frac{x - np}{\sqrt{npq}}$, $q = 1-p$ is approximately $N(0,1)$, i.e.,

If $U = \frac{x - np}{\sqrt{npq}}$ then $U \approx N(0,1)$

Choose $u_\varepsilon > 0$ such that

$$P(-u_\varepsilon < U < u_\varepsilon) = 1 - \varepsilon \quad \text{———— (1)}$$



$$\Rightarrow P(|U| < u_\varepsilon) = 1 - \varepsilon$$

$$\Rightarrow 1 - P(|U| \geq u_\varepsilon) = 1 - \varepsilon$$

$$\Rightarrow P(|U| \geq u_\varepsilon) = \varepsilon$$

$$\Rightarrow P(|U| > u_\varepsilon) = \varepsilon$$

$$\Rightarrow P(U < -u_\varepsilon) + P(U > u_\varepsilon) = \varepsilon$$

$$\Rightarrow P(U > u_\varepsilon) + P(U > u_\varepsilon) = \varepsilon$$

$$\Rightarrow 2P(U > u_\varepsilon) = \varepsilon$$

$$\Rightarrow P(U > u_\varepsilon) = \frac{\varepsilon}{2} \quad \text{———— (2)}$$

Therefore, one can obtain u_ε from equation (2).

Again from ① we get

$$P(-u_\varepsilon < U < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|U| < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|U| - u_\varepsilon < 0) = 1-\varepsilon$$

$$\Rightarrow P((|U| + u_\varepsilon)(|U| - u_\varepsilon) < 0) = 1-\varepsilon \quad [\text{as } |U| + u_\varepsilon > 0]$$

$$\Rightarrow P(|U|^2 - u_\varepsilon^2 < 0) = 1-\varepsilon$$

$$\Rightarrow P(U^2 - u_\varepsilon^2 < 0) = 1-\varepsilon$$

$$\Rightarrow P\left(\left(\frac{x-np}{\sqrt{npq}}\right)^2 - u_\varepsilon^2 < 0\right) = 1-\varepsilon$$

$$\Rightarrow P((x-np)^2 - npq\, u_\varepsilon^2 < 0) = 1-\varepsilon$$

$$\Rightarrow P(x^2 - 2xnp + n^2p^2 - np(1-p)u_\varepsilon^2 < 0) = 1-\varepsilon$$

$$\Rightarrow P(x^2 - 2xnp + n^2p^2 - np\,u_\varepsilon^2 + np^2u_\varepsilon^2 < 0) = 1-\varepsilon$$

$$\Rightarrow P(np^2(n+u_\varepsilon^2) - 2xnp - np\,u_\varepsilon^2 + x^2 < 0) = 1-\varepsilon$$

$$\Rightarrow P(np^2(n+u_\varepsilon^2) - np(2x+u_\varepsilon^2) + x^2 < 0) = 1-\varepsilon$$

$$\Rightarrow P\left(p^2 - \frac{np(2x+u_\varepsilon^2)}{n(n+u_\varepsilon^2)} + \frac{x^2}{n(n+u_\varepsilon^2)} < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(p^2 - p\,\frac{2x+u_\varepsilon^2}{n+u_\varepsilon^2} + \frac{x^2}{n(n+u_\varepsilon^2)} < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(p^2 - 2\cdot p\cdot\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)} + \left\{\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right\}^2 - \left\{\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right\}^2 + \frac{x^2}{n(n+u_\varepsilon^2)} < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\left\{p - \frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right\}^2 - \left[\sqrt{\left\{\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right\}^2 - \frac{x^2}{n(n+u_\varepsilon^2)}}\,\right]^2 < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\left|p - \frac{2x + u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right|^2 - \delta^2 < 0\right) = 1-\varepsilon^2 \quad \text{——} \; (3)$$

where

$$\delta = \sqrt{\left\{\frac{2x + u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right\}^2 - \frac{x^2}{n(n+u_\varepsilon^2)}}$$

$$= \sqrt{\frac{(2x+u_\varepsilon^2)^2}{4(n+u_\varepsilon^2)^2} - \frac{n\,x^2}{n^2(n+u_\varepsilon^2)}}$$

$$= \sqrt{\frac{n^2(2x+u_\varepsilon^2)^2 - 4nx^2(n+u_\varepsilon^2)}{4n^2(n+u_\varepsilon^2)^2}}$$

$$= \frac{\sqrt{n^2(2x+u_\varepsilon^2)^2 - 4nx^2(n+u_\varepsilon^2)}}{2n(n+u_\varepsilon^2)}$$

$$= \frac{\sqrt{n^2(4x^2+4xu_\varepsilon^2+u_\varepsilon^4) - 4n^2x^2 - 4nx^2u_\varepsilon^2}}{2n(n+u_\varepsilon^2)}$$

$$= \frac{\sqrt{4n^2x^2 + 4n^2xu_\varepsilon^2 + n^2u_\varepsilon^4 - 4n^2x^2 - 4nx^2u_\varepsilon^2}}{2n(n+u_\varepsilon^2)}$$

$$= \frac{\sqrt{4nxu_\varepsilon^2(n-x) + n^2u_\varepsilon^4}}{2n(n+u_\varepsilon^2)}$$

$$= \frac{\sqrt{4nxu_\varepsilon^2(n-x) + n^2u_\varepsilon^4}}{2n(n+u_\varepsilon^2)}$$

As $X = B(n,p)$, $x = i$, $i = 0(1)n$

$\Rightarrow x < n \Rightarrow n - x > 0$

$\therefore$ δ is well defined as a strictly positive real number.

From ③ we get

$$P\left(\left|p-\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right|^2-\delta^2<0\right)=1-\varepsilon$$

$$\Rightarrow P\left(\underbrace{\left[\left|p-\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right|+\delta\right]}_{>0}\left[\left|p-\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right|-\delta\right]<0\right)=1-\varepsilon$$

$$\Rightarrow P\left(\left|p-\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right|-\delta<0\right)=1-\varepsilon$$

$$\Rightarrow P\left(\left|p-\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}\right|<\delta\right)=1-\varepsilon$$

$$\Rightarrow P\left(-\delta<p-\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}<\delta\right)=1-\varepsilon$$

$$\Rightarrow P\left(\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}-\delta<p<\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}+\delta\right)=1-\varepsilon$$

$$\Rightarrow P\left(A<p<B\right)=1-\varepsilon$$

where

$$A=\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}-\delta=\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}-\frac{\sqrt{4nxu_\varepsilon^2(n-x)+n^2u_\varepsilon^4}}{2n(n+u_\varepsilon^2)}$$

$$B=\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}+\delta=\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}+\frac{\sqrt{4nxu_\varepsilon^2(n-x)+n^2u_\varepsilon^4}}{2n(n+u_\varepsilon^2)}$$

$\therefore$ The approximate confidence interval of p is (a,b) where

$$a=\frac{2x+u_\varepsilon^2}{2(n+u_\varepsilon^2)}-\frac{\sqrt{4nxu_\varepsilon^2(n-x)+n^2u_\varepsilon^4}}{2n(n+u_\varepsilon^2)},$$

$$b = \frac{2x + u_\varepsilon^2}{2(n+u_\varepsilon^2)} + \frac{\sqrt{4nxu_\varepsilon^2(n-x) + n^2 u_\varepsilon^4}}{2n(n+u_\varepsilon^2)},$$

and x is the observed value of X.

Problem 5: Show that the approximate confidence interval of p (at the confidence coefficient $1-\varepsilon$) for Binomial (n, p) population is

$$\left(\frac{x}{n} - u_\varepsilon \sqrt{\frac{x(n-x)}{n^3}},\ \frac{x}{n} + u_\varepsilon \sqrt{\frac{x(n-x)}{n^3}} \right),$$

where u_ε is given by

$$P(U > u_\varepsilon) = \frac{\varepsilon}{2}, \quad U \equiv N(0,1)$$

and it is assumed that p is replaced by its maximum likelihood estimate $\hat{p} = \frac{x}{n}$ to calculate the standard deviation of $X = B(n, p)$. Here x is the observed value of the samble.

Answer: The mean of $X = np$

The standard deviation of $X = \sqrt{npq}$

$$= \sqrt{np(1-p)} = \sqrt{n\hat{p}(1-\hat{p})} = \sqrt{n\frac{x}{n}\left(1-\frac{x}{n}\right)}$$

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

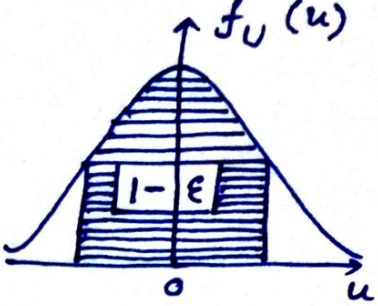
$\therefore$ For large n, $\dfrac{X-np}{\sqrt{npq}} \approx N(0,1)$

$$\Rightarrow \frac{X-np}{\sqrt{\frac{x(n-x)}{n}}} \approx N(0,1)$$

$$\Rightarrow U = \frac{X-np}{\sqrt{\frac{x(n-x)}{n}}} \approx N(0,1)$$

choose $u_\varepsilon > 0$ such that

$$P(-u_\varepsilon < U < u_\varepsilon) = 1-\varepsilon \quad \text{——— (1)}$$



$$\Rightarrow P(|U| < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|U| < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow 1 - P(|U| \geq u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|U| \geq u_\varepsilon) = \varepsilon$$

$$\Rightarrow 2P(U > u_\varepsilon) = \varepsilon$$

$$\Rightarrow P(U > u_\varepsilon) = \varepsilon/2 \quad \text{——— (2)}$$

Again from (1) we get

$$P(-u_\varepsilon < U < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|U| < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P\left(\left|\frac{x-np}{\sqrt{\frac{x(n-x)}{n}}}\right| < u_\varepsilon\right) = 1-\varepsilon$$

$$\Rightarrow P\left(|x-np| < u_\varepsilon\sqrt{\frac{x(n-x)}{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(|np-x| < u_\varepsilon\sqrt{\frac{x(n-x)}{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(-u_\varepsilon\sqrt{\frac{x(n-x)}{n}} < np-x < u_\varepsilon\sqrt{\frac{x(n-x)}{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(x - u_\varepsilon\sqrt{\frac{x(n-x)}{n}} < np < x + u_\varepsilon\sqrt{\frac{x(n-x)}{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\frac{x}{n} - \frac{u_\varepsilon}{n}\sqrt{\frac{x(n-x)}{n}} < p < \frac{x}{n} + \frac{u_\varepsilon}{n}\sqrt{\frac{x(n-x)}{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\frac{x}{n} - u_\varepsilon\sqrt{\frac{x(n-x)}{n^3}} < p < \frac{x}{n} + u_\varepsilon\sqrt{\frac{x(n-x)}{n^3}}\right) = 1-\varepsilon$$

$$\Rightarrow P(A < p < B) = 1-\varepsilon$$

$$A = \frac{x}{n} - u_\varepsilon\sqrt{\frac{x(n-x)}{n^3}}, \quad B = \frac{x}{n} + u_\varepsilon\sqrt{\frac{x(n-x)}{n^3}}$$

Therefore the confidence interval of p at the confidence level $1-\varepsilon$ is (a, b), where

$$a = \frac{x}{n} - u_\varepsilon\sqrt{\frac{x(n-x)}{n^3}}, \quad b = \frac{x}{n} + u_\varepsilon\sqrt{\frac{x(n-x)}{n^3}}$$

<u>Problem 6</u>: Make a Correspondence between problem ④ & problem ⑤

<u>Answer</u>: From problem ④ we get

$$a, b = \frac{2x + u_\varepsilon^2}{2(n + u_\varepsilon^2)} \mp \frac{\sqrt{4nx\,u_\varepsilon^2(n-x) + n^2 u_\varepsilon^4}}{2n(n+u_\varepsilon^2)}$$

where
$$P(U > u_\varepsilon) = \varepsilon/2, \quad U = N(0,1)$$

$\therefore$ u_ε depends only on ε, and consequently for fixed ε, u_ε is a constant.

Now
$$\frac{2x + u_\varepsilon^2}{2(n+u_\varepsilon^2)} = \frac{2x + u_\varepsilon^2}{2n(1+u_\varepsilon^2/n)} = \frac{2x+u_\varepsilon^2}{2n}\left(1+\frac{u_\varepsilon^2}{n}\right)^{-1}$$

$$= \frac{2x + u_\varepsilon^2}{2n}\left(1 - \frac{u_\varepsilon^2}{n} + \cdots\right)$$

$$= \left(\frac{x}{n} + \underset{\downarrow\, 0}{\frac{u_\varepsilon^2}{2n}}\right)\left(1 - \underset{\downarrow\, 0}{\frac{u_\varepsilon^2}{n}} + \cdots\right) = \frac{x}{n}$$

Again $\dfrac{\sqrt{4nx\,u_\varepsilon^2(n-x)+n^2u_\varepsilon^4}}{2n(n+u_\varepsilon^2)}$

$$= u_\varepsilon \frac{\sqrt{4nx(n-x)+n^2u_\varepsilon^2}}{2n(n+u_\varepsilon^2)}$$

$$= u_\varepsilon \sqrt{\frac{4nx(n-x)+n^2u_\varepsilon^2}{4n^2(n+u_\varepsilon^2)^2}}$$

$$= u_\varepsilon \sqrt{\frac{4nx(n-x)+n^2u_\varepsilon^2}{4n^4(1+u_\varepsilon^2/n)^2}}$$

$$= u_\varepsilon \sqrt{\frac{x(n-x)}{n^3}\underbrace{\left(1+\frac{u_\varepsilon^2}{n}\right)^{-2}}_{\to 1}+\frac{u_\varepsilon^2}{4n^2}\underbrace{\left(1+\frac{u_\varepsilon^2}{n}\right)^{-2}}_{\to 1}}$$

$$= u_\varepsilon \sqrt{\frac{x(n-x)}{n^3}+\underbrace{\frac{u_\varepsilon^2}{4n^2}}_{\to 0}}$$

$$= u_\varepsilon \sqrt{\frac{x(n-x)}{n^3}}$$

$$\therefore\ a, b = \frac{x}{n} \mp u_\varepsilon \sqrt{\frac{x(n-x)}{n^3}}$$

$\therefore$ An approximation of a, b upto a order of $\frac{1}{n}$ with respect to a constant results the answer of problem 5.

<u>Problem 7</u>: Show that the approximate confidence limits for large samples for the parameter μ of a Poisson population having confidence coefficient $1-\varepsilon$ are the roots of the quadratic equation in μ:

$$n(\bar{x}-\mu)^2 = u_\varepsilon \mu,$$

where $\bar{x}$ is the sample mean for a sample of size n and u_ε is given by

$P(U > u_\varepsilon) = \frac{\varepsilon}{2}$ with $U = N(0,1)$

<u>Answer</u>: Let $x_1, x_2, \ldots, x_n$ be a sample of size n drawn from the population of $X = P(\mu)$. If $X_1, X_2, \ldots, X_n$ be the random variables corresponding to the sample values $x_1, x_2, \ldots, x_n$ respectively, then $X_1, X_2, \ldots, X_n$ are mutually independent and

$X_i = P(\mu) \quad \forall\ i = 1(1)n$

For large sample, $\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$ can be assumed as a normal variate with

mean $= E(X)$ and variance $=$ Variance $(X)/n$

$$\therefore \bar{X} \approx N\left(E(X), \frac{\sqrt{Var(X)}}{\sqrt{n}}\right)$$

We know that for Poisson μ distribution

$E(X) = \mu$ and $Var(X) = \mu$

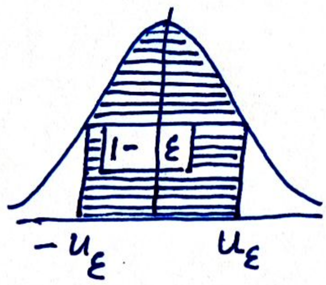
$\therefore \bar{X} \approx N\left(\mu, \sqrt{\frac{\mu}{n}}\right)$

$$\Rightarrow \frac{\bar{X}-\mu}{\sqrt{\frac{\mu}{n}}} \approx N(0,1)$$

$$\Rightarrow \sqrt{\frac{n}{\mu}}\,(\bar{X}-\mu) \approx N(0,1)$$

$$\Rightarrow U = \sqrt{\frac{n}{\mu}}\,(\bar{X}-\mu) \approx N(0,1)$$

Choose a strictly positive real number u_ε such that



$$P(-u_\varepsilon < U < u_\varepsilon) = 1-\varepsilon \quad \text{——— (1)}$$

$$\Rightarrow P(|U| < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow 1 - P(|U| \geq u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|U| \geq u_\varepsilon) = \varepsilon$$

$$\Rightarrow 2P(U \geq u_\varepsilon) = \varepsilon$$

$$\Rightarrow P(U \geq u_\varepsilon) = \varepsilon/2$$

$$\Rightarrow P(U > u_\varepsilon) = \varepsilon/2 \quad \text{——— (2)}$$

This equation determines the value of u_ε for given ε.

Again from (1) we get

$$P(-u_\varepsilon < U < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|U| < u_\varepsilon) = 1-\varepsilon$$

$$\Rightarrow P(|U| - u_\varepsilon < 0) = 1-\varepsilon$$

$$\Rightarrow P((|U| + u_\varepsilon)(|U| - u_\varepsilon) < 0) = 1-\varepsilon$$

$$\Rightarrow P(|U|^2 - u_\varepsilon^2 < 0) = 1-\varepsilon$$

$$\Rightarrow P(U^2 - u_\varepsilon^2 < 0) = 1-\varepsilon$$

$$\Rightarrow P\left(\frac{n}{\mu}(\bar{x}-\mu)^2 - u_\varepsilon^2 < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left((\bar{x}-\mu)^2 - \frac{\mu u_\varepsilon^2}{n} < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\bar{x}^2 - 2\mu\bar{x} + \mu^2 - \frac{\mu u_\varepsilon^2}{n} < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\mu^2 - 2\mu\left(\bar{x} + \frac{u_\varepsilon^2}{2n}\right) + \bar{x}^2 < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\left[\mu - \left(\bar{x} + \frac{u_\varepsilon^2}{2n}\right)\right]^2 - \left(\bar{x} + \frac{u_\varepsilon^2}{2n}\right)^2 + \bar{x}^2 < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\left[\mu - \left(\bar{x} + \frac{u_\varepsilon^2}{2n}\right)\right]^2 - \bar{x}^2 - \frac{u_\varepsilon^4}{4n^2} - \bar{x}\frac{u_\varepsilon^2}{n} + \bar{x}^2 < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\left[\mu - \left(\bar{x} + \frac{u_\varepsilon^2}{2n}\right)\right]^2 - \left(\frac{u_\varepsilon^4}{4n^2} + \bar{x}\frac{u_\varepsilon^2}{n}\right) < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\left|\mu - \left(\bar{x} + \frac{u_\varepsilon^2}{2n}\right)\right|^2 - \left(\sqrt{\frac{u_\varepsilon^4}{4n^2} + \bar{x}\frac{u_\varepsilon^2}{n}}\right)^2 < 0\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\left|\mu - \left(\bar{x} + \frac{u_\varepsilon^2}{2n}\right)\right| < \sqrt{\frac{u_\varepsilon^4}{4n^2} + \bar{x}\frac{u_\varepsilon^2}{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P\left(\bar{x} + \frac{u_\varepsilon^2}{2n} - \sqrt{\frac{u_\varepsilon^4}{4n^2} + \bar{x}\frac{u_\varepsilon^2}{n}} < \mu < \bar{x} + \frac{u_\varepsilon^2}{2n} + \sqrt{\frac{u_\varepsilon^4}{4n^2} + \bar{x}\frac{u_\varepsilon^2}{n}}\right) = 1-\varepsilon$$

$$\Rightarrow P(A < \mu < B) = 1-\varepsilon$$

where

$$A = \bar{x} + \frac{u_\varepsilon^2}{2n} - \sqrt{\frac{u_\varepsilon^4}{4n^2} + \bar{x}\,\frac{u_\varepsilon^2}{n}}$$

$$B = \bar{x} + \frac{u_\varepsilon^2}{2n} + \sqrt{\frac{u_\varepsilon^4}{4n^2} + \bar{x}\,\frac{u_\varepsilon^2}{n}}$$

Therefore the confidence interval for μ is (a, b),

where

$$a = \bar{x} + \frac{u_\varepsilon^2}{2n} - \sqrt{\frac{u_\varepsilon^4}{4n^2} + \bar{x}\,\frac{u_\varepsilon^2}{n}}$$

$$b = \bar{x} + \frac{u_\varepsilon^2}{2n} + \sqrt{\frac{u_\varepsilon^4}{4n^2} + \bar{x}\,\frac{u_\varepsilon^2}{n}}$$

$$\therefore\ a + b = 2\left(\bar{x} + \frac{u_\varepsilon^2}{2n}\right)$$

$$ab = \left(\bar{x} + \frac{u_\varepsilon^2}{2n}\right)^2 - \left(\frac{u_\varepsilon^4}{4n^2} + \bar{x}\,\frac{u_\varepsilon^2}{n}\right)$$

$$= \bar{x}^2$$

$\therefore$ a, b are the roots of the quadratic equation in μ:

$$\mu^2 - (a+b)\mu + ab = 0$$

$$\Rightarrow \mu^2 - 2\left(\bar{x} + \frac{u_\varepsilon^2}{2n}\right)\mu + \bar{x}^2 = 0$$

$$\Rightarrow \bar{x}^2 - 2\bar{x}\mu + \mu^2 - \frac{u_\varepsilon^2}{n}\mu = 0$$

$$\Rightarrow (\bar{x} - \mu)^2 = \frac{u_\varepsilon^2}{n}\mu$$

$$\Rightarrow n(\bar{x} - \mu)^2 = u_\varepsilon^2 \mu$$

Testing of Hypothesis

The functional form of a probability distribution of a population is generally unknown or the functional form of the probability distribution of a population may be known but it contains a number of unknown parameters. A statistical hypothesis is any assumption regarding the probability distribution of population on the basis of a sample of any size drawn from the population.

Here we assume that the functional form of the probability distribution of population is known but it contains a number of unknown parameters. Let $F(x)$ be the probability distribution function of a population corresponding to a random variable X. Let $F(x)$ contains a number of unknown parameters $\theta_1, \theta_2, \ldots, \theta_k$. As the functional form of $F(x)$ is known, the statistical hypothesis, for this case, is an assumption regarding the k-dimensional point $\Theta \in P_K$, where

① $\Theta = (\theta_1, \theta_2, \dots, \theta_K)$ and

② P_K is the k-dimensional parametric space or k-dimensional parameter space.

⊕ As there are many restrictions on the unknown parameters $\theta_1, \theta_2, \dots, \theta_K$, $P_K \neq E_K$. For example, consider a normal $N(m, \sigma)$ population, where the values of the parameters m and σ are unknown.

As $\sigma > 0$

$$P_2 = \{(m, \sigma) : \sigma > 0\} \subseteq E_2 \quad \text{and} \quad P_2 \neq E_2$$

⊛ A statistical hypothesis on the parameters is any assumption of the following form:

$H_0 : \Theta \in \omega$ and $\omega \subset P_K$ with $\omega \neq P_K$

where the functional form of the distribution of the population is known.

- if ω contains a single point $\Theta_0 = (\theta_{10}, \theta_{20}, \ldots, \theta_{k0})$ then

$H_0 : \Theta \in \omega = \{(\theta_{10}, \theta_{20}, \ldots, \theta_{k0})\}$

$\Downarrow$

$H_0 : \Theta = \Theta_0 \Rightarrow H_0 : \theta_1 = \theta_{10}, \ \theta_2 = \theta_{20}, \ldots, \theta_k = \theta_{k0}$

and in this case H_0 is known as simple hypothesis

- if ω contains more than one point of P_k then H_0 is known as composite hypothesis.

- Example : Consider a normal $N(m, \sigma)$ population

(a) $H_0 : m = 2, \ \sigma = 0.01 \longrightarrow$ simple hypothesis

(b) $H_0 : m = 2$, σ is not specified $\Rightarrow$

$H_0 : m = 2, \ \sigma > 0 \Rightarrow H_0 : (m, \sigma) \in \{(2, \sigma) : \sigma > 0\} = \omega$

$\Rightarrow$ ω contains more than one points

$\Rightarrow$ H_0 is composite hypothesis.

(c) $H_0 : 2 < m < 3, \ \sigma > 0 \longrightarrow$ composite hypothesis

(d) $H_0 : m = \sigma^2 \longrightarrow$ composite hypothesis

(e) $H_0 : \sigma = m^2 \longrightarrow$ composite hypothesis

Null hypothesis: Sometimes, we want to make a hypothesis wishing it to be rejected by the test. This hypothesis is known as Null hypothesis and generally denoted by H_0.

Alternative hypothesis: The alternative hypothesis against the hypothesis

$H_0 : \Theta \in \omega$

is generally denoted by H_1 and defined by

$H_1 : \Theta \in \bar{\omega} = P_K - \omega$

Note: Suppose we want to test the hypothesis

$H_0 : \Theta \in \omega \subseteq P_K$

against no alternative.

Actully, here we want to test H_0 against the alternative

$H_1 : \Theta \notin \omega$ or $H_1 : \Theta \in P_K - \omega$

<u>Example</u>: Consider a normal $N(m, \sigma)$ population

Suppose we want to test the hypothesis

$H_0 : m = m_0$

against no alternative

$\Downarrow$

We want to test H_0 against the alternative hypothesis H_1 given by

$H_1 : m \neq m_0$

General form of a test of hypothesis:

Let us consider a sample of size n : $x_1, x_2, \ldots, x_n$ drawn from a population of the random variable X. The functional form of the probability distribution of X is known but it contains a number of unknown parameters $\theta_1, \theta_2, \ldots, \theta_k$. Let us consider a hypothesis

$H_0 : \Theta = (\theta_1, \theta_2, \ldots, \theta_k) \in \omega \subseteq R_k$

Our aim is to test the hypothesis H_0 on the basis of the sample $x_1, x_2, \dots, x_n$ drawn from the population of X.

① To test the hypothesis H_0, we want to construct a region W in sample space such that

(I) if $\underset{\sim}{x}=(x_1, x_2, \dots, x_n) \in W$, H_0 is rejected

(II) if $\underset{\sim}{x}=(x_1, x_2, \dots, x_n) \in \overline{W}$, H_0 is accepted

In this case, W is known as rejection region whereas $\overline{W}$ is known as acceptance region

② There are two types of errors in choosing W from sample space

(A) Type I error : H_0 is true but $(x_1, x_2, \dots, x_n) \in W$

$\Rightarrow$ H_0 is true but the observed value of the sample falls in rejection region and consequently we reject H_0

$\Rightarrow$ H_0 is true but H_0 is rejected by the test

$\Rightarrow x \in W$ but H_0 is true

$\Rightarrow x \in W : \Theta \in \omega$

(B) <u>Type II error</u> : H_0 is false but $(x_1, x_2, \ldots, x_n) \in \overline{W}$

$\Rightarrow$ H_0 is false but the observed value of the sample falls in the acceptance region and consequently we accept H_0

$\Rightarrow$ H_0 is false but H_0 is accepted by the test

$\Rightarrow$ $x \in \overline{W}$ but H_0 is false

$\Rightarrow$ $x \in \overline{W} : \Theta \in \overline{\omega}$

For constructing good test, we generally want to minimize the probabilities of both type of errors as much as possible. But it is found that if the probability of type I error decreases then the probability of type II error automatically increases and it is not possible to make the probabilities of both type of errors simultaneously small.

Hence in order to obtain best test, the general prescription is to fix the probability of type I error at a given significance level ε, $0<\varepsilon<1$,

i.e., $P(x \in W \mid \Theta \in \omega) = \varepsilon$

and then we minimize the probility of type 2 error and the region W of sample space so obtained is known as best critical region or best rejection region.

<u>Power of the test</u>: Power of the test is the probability of rejecting a hypothesis when it is actually false. It is generally denoted by $\beta(W)$

<u>Theorem</u>: $\beta(W) + P(\text{type II error}) = 1$

<u>Proof</u>: $P(\text{type II error}) = P(x \in \bar{W} \mid \Theta \in \bar{\omega})$

$= 1 - P(x \in W \mid \Theta \in \bar{\omega})$

$= 1 -$ Probability of rejecting the hypothesis when it is false

$= 1 - \beta(W) \Rightarrow \boxed{\beta(W) + P(\text{type II error}) = 1}$

✓ Likelihood ratio testing

Suppose the functional form of the distribution of the population is known but it contains a number of unknown parameters. Let us consider a hypothesis Ho such that the number of unknown parameters under Ho is reduced.

Let us consider a sample of size n: $x_1, x_2, \ldots, x_n$ drawn from the population of X. Then the likelihood function for the given sample values can be written as

$$L(x;\Theta) = L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k)$$

where $\theta_1, \theta_2, \ldots, \theta_k$ are the unknown parameters.

Let

$$L(x;\Theta') = L(x;\Theta|_{H_0})$$

= Likelihood function under Ho for the same sample

∴ the number of unknown parameters in Θ' $<$ the number of unknown parameters in $\Theta = k$

Let $\hat{\Theta}$ be the maximum likelihood estimate of Θ for the likelihood function $L(x; \Theta)$

$$\Rightarrow L(x; \hat{\Theta}) = \max L(x; \Theta)$$

Again, let $\hat{\Theta}'$ be the maximum likelihood estimate of Θ' for the likelihood function $L(x; \Theta')$

$$\therefore L(x; \hat{\Theta}') = \max L(x; \Theta')$$

$$\therefore \quad 0 \leq L(x; \hat{\Theta}') \leq L(x; \hat{\Theta})$$

Maximum is taken over a sub space of P_K

Maximum is taken over P_K

$$\Rightarrow 0 \leq \frac{L(x; \hat{\Theta}')}{L(x; \hat{\Theta})} \leq 1$$

$$\Rightarrow 0 \leq \lambda \leq 1 \quad \text{——— (1)}$$

where

$$\lambda = \frac{L(x; \hat{\Theta}')}{L(x; \hat{\Theta})}$$

As likelihood estimate of any unknown parameter function of sample values,

① $L(x; \hat{\Theta}')$ is free from any unknown parameters

② $L(x; \hat{\Theta})$ is also free from any unknown parameters

and therefore

$\lambda = \lambda(x)$ is free from any unknown parameters and this statistic is known as <u>likelihood statistic</u> under the hypothesis H_0.

Again inequality ① shows that the spectrum of the corresponding variate λ is the closed interval $[0, 1]$.

Ⓐ If H_0 is true,

$$L(x; \Theta') = L(x; \Theta)$$

$$\left.\begin{aligned} L(x; \hat{\Theta}') &\simeq L(x; \Theta') \\ L(x; \hat{\Theta}) &\simeq L(x; \Theta) \end{aligned}\right\}$$ As the maximum likelihood estimates are known to be good estimates of the unknown parameters

$$\Downarrow$$

$$L(x; \hat{\Theta}') \simeq L(x; \hat{\Theta})$$

$$\Rightarrow \frac{L(x;\hat{\theta}')}{L(x;\hat{\theta})} \simeq 1$$

$$\Rightarrow \lambda \simeq 1$$

$\Rightarrow$ If the observed value of $\lambda(x)$ is close to 1, then we believe that H_0 is true and we accept H_0

On the otherhand, if the observed value of λ is close to zero then $L(x;\hat{\theta}') << L(x;\hat{\theta})$

$\Rightarrow H_0$ is false

Hence for testing H_0, we can take λ as the statistic of the test for which the critical region will be $(0, \lambda_\varepsilon)$ and the probability of type I error is given by

$$P(0<\lambda<\lambda_\varepsilon \mid H_0) = \int_0^{\lambda_\varepsilon} f_\lambda(\lambda)\, d\lambda = \varepsilon$$

$\uparrow$

This equation uniquely determine λ_ε for given value of ε.

If the observed value of $\lambda \in (0, \lambda_\varepsilon)$ then we reject H_0, otherwise we accept H_0.

Problem 1 : Let $f(x)$ be the probability density function of the population of a random variable X, where

$$f(x) = \begin{cases} \frac{1}{\theta} , & 0 \le x \le \theta \\ 0 , & \text{elsewhere} \end{cases}$$

and θ is the only unknown parameter. We are testing the null hypothesis $H_0 : \theta = 1$ against $H_1 : \theta = 2$ by means of a single observed value x of X. If

(a) $W = \{x : 0.5 \le x\}$ is the critical region then find the probability of type I error and the power of the test

(b) $W = \{x : 1 \le x \le 1.5\}$ is the critical region then find the probability of type I error and the power of the test.

Answer : (a) Probability of type I error

$= P(x \in W \text{ but } H_0 \text{ is true})$

$= P(0.5 \le x \le \theta : H_0 \text{ is true} \Rightarrow \theta = 1)$

$= P(0.5 \le x \le \theta : \theta = 1) = P(0.5 \le x \le 1) = \int_{0.5}^{1} f(x)\,dx = \int_{0.5}^{1} 1\,dx \; [\because \theta = 1]$

$= 0.5$

Power of the test = Probability of rejecting H_0 when it is false

$= P(x \in W : H_0 \text{ is false})$

$= P(x \in W : H_1 \text{ is true})$

$= P(0.5 \leq x \leq \theta : \theta = 2)$

$= P(0.5 \leq x \leq 2)$ $\boxed{\text{with } \theta = 2}$

$$= \int_{0.5}^{2} f_x \, dx = \int_{0.5}^{2} \frac{1}{2} \, dx = \frac{1}{2}(2 - 0.5) = \frac{1}{2} \times 1.5 = 0.75$$

$\Rightarrow \beta$ = power of the test $= 0.75$

(b) Here $W = \{x : 1 \leq x \leq 1.5\}$

Probability of type I error

= Probability of rejecting H_0 when it is true)

$= P(x \in W : H_0 = 1)$

$= P(1 \leq x \leq 1.5)$

$$= \int_{1}^{1.5} f(x) \, dx$$

$$= \int_{1}^{1.5} 0 \cdot dx = 0$$

$$f(x) = \begin{cases} \frac{1}{\theta} & 0 \leq x \leq \theta \\ 0 & \text{elsewhere} \end{cases}$$

$\Downarrow$

$$f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{cases} \quad \text{for } \theta = 1$$

Power of the test

= Probability of rejecting H_0 when it is false

$= P(x \in W : H_0$ is false$)$ $\qquad f(x) = \begin{cases} \frac{1}{\theta}, & 0 \le x \le \theta \\ 0, & \text{otherwise} \end{cases}$

$= P(x \in W : H_1$ is true$)$ $\qquad = \begin{cases} \frac{1}{2}, & 0 \le x \le 2 \\ 0, & \text{otherwise} \end{cases}$

$= P(1 \le x \le 1.5 : \theta = 2)$

$$= \int_1^{1.5} f(x)\,dx = \int_1^{1.5} \frac{1}{2}\,dx = \frac{1}{2}(1.5-1) = \frac{1}{2} \times 0.5 = 0.5 \times 0.5$$

$= 0.25$

Problem 2: If $W = \{x : x \ge 1\}$ is the critical region for testing $H_0 : \theta = 2$ against the alternative $H_1 : \theta = 1$ on the basis of the single observation x from the population having probability density function

$f(x) = \theta e^{-\theta x}, \quad 0 \le x < \infty$

for the unknown parameter θ, obtain the probability of type I error and the power of the test.

Problem 3: Let $X = N(\mu, 2)$ where μ is a unknown parameter, i.e., $E(X) = \mu$ and $Var(X) = 2^2 = 4$. We are testing the hypothesis $H_0: \mu = -1$ against the alternative hypothesis $H_1: \mu = 1$ on the basis of a sample $x_1, x_2, \ldots, x_{10}$ drawn from the population of X. If the critical region

$$W = \{(x_1, x_2, \ldots, x_{10}) : x_1 + 2x_2 + \cdots + 10x_{10} \geq 0\},$$

then find the power of the test. Given that

$$\frac{1}{\sqrt{2\pi}} \int_0^{1.6} e^{-\frac{u^2}{2}} du = 0.4192$$

Answer: Let $u = x_1 + 2x_2 + \cdots + 10x_{10}$

$\Rightarrow U = X_1 + 2X_2 + \cdots + 10X_{10}$

$$E(U) = E(X_1) + 2E(X_2) + \cdots + 10E(X_{10}) = \mu + 2\mu + \cdots + 10\mu = 55\mu$$

$$Var(U) = 1^2 Var(X_1) + 2^2 Var(X_2) + \cdots + 10^2 Var(X_{10})$$

$$= (1^2 + 2^2 + \cdots + 10^2) Var(X) = \frac{10(10+1)(2 \times 10 + 1)}{6} Var(X)$$

$$= \frac{\cancel{10}^{5} \times 11 \times \cancel{21}^{7}}{\cancel{6}_{3}} Var(X) = 385 \, Var(X) = 385 \times 4 = 1540$$

$$U = N(55\mu, 1540) \Rightarrow \frac{U - 55\mu}{\sqrt{1540}} = N(0,1)$$

$$\Rightarrow Z = \frac{U - 55\mu}{\sqrt{1540}} = N(0,1)$$

Power of the test

$= $ Probability of rejecting H_0 when it is false

$= P(x \in W : H_0 \text{ is false}), \quad x = (x_1, x_2, \ldots, x_n)$

$= P(x \in W : H_1 \text{ is true})$

$= P(x \in W : \mu = 1)$

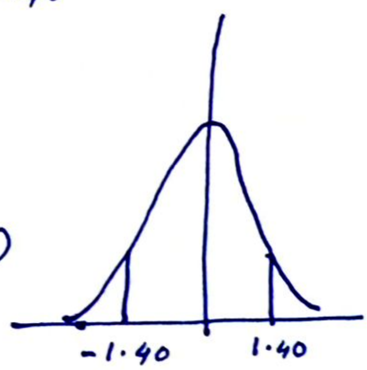
$= P(u \geq 0 : \mu = 1) \qquad z = \dfrac{u - 55\mu}{\sqrt{1540}} \Rightarrow 55\mu + \sqrt{1540}\,z = u$

$= P(55\mu + \sqrt{1540}\,z \geq 0 : \mu = 1)$

$= P(55 + \sqrt{1540}\,z \geq 0)$

$= P\left(z \geq -\dfrac{55}{\sqrt{1540}}\right) \qquad \dfrac{55}{\sqrt{1540}} \approx 1.40$

$= P(z \geq -1.40)$



$= P(-1.40 \leq z \leq 0) + P(0 \leq z < \infty)$

$= P(0 \leq z \leq 1.40) + 0.5$

$= 0.5 + 0.4192 = 0.9192$