Standard Error. The standard deviation of a sampling distribution of a sample statistic is called the *standard error* (abbreviated to *S.E.*) *of the statistic*. The standard error thus measures the dispersion of the values of a statistic, that might be computed from all possible samples, whereas the standard deviation of a population (or sample) measures the dispersion of the values of the population (sample) units about the population (sample) mean.

14.4.1. Sampling Distribution of the Mean. The sampling distribution of the mean is the probability distribution or the relative frequency distribution of the means $\bar{X}$ of all possible random samples of the same size that *could* be selected from a given population. The mean of this distribution is represented by $\mu_{\bar{x}}$ and the standard deviation, which is called the *standard error of the mean*, by $\sigma_{\bar{x}}$ or *S.E.* ($\bar{X}$). The value $\sigma_{\bar{x}}$ indicates the spread in the distribution of all possible sample means.

The sampling distribution of $\bar{X}$ has the following properties.

(i) The mean of the sampling distribution of the mean (equivalently, the mean of all possible sample means) is equal to the population mean, that is $\mu_{\bar{x}} = \mu$, regardless of whether sampling is done with replacement or without replacement.

Proof. Let us first consider sampling *without replacement* from a finite population of size N. The number of distinct simple random samples of size n that can be selected without replacement from a population of size N is $\binom{N}{n} = k$, say. Let $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_i, \ldots, \bar{X}_k$ be the means of $k = \binom{N}{n}$ possible random samples of size n, where $\bar{X}_i$ is the mean of the *ith* sample. Then the mean of the sampling distribution of $\bar{X}$ (equivalently, the mean of all possible sample means), denoted by $\mu_{\bar{x}}$, is

$$\mu_{\bar{x}} = \frac{\bar{X}_1 + \bar{X}_2 + \ldots + \bar{X}_k}{k}$$

$$= \frac{1}{k}\left[\left(\frac{X_1+X_2+\ldots}{n}\right) + \left(\frac{X_1+X_3+\ldots}{n}\right) + \ldots + \left(\frac{X_2+X_3+\ldots}{n}\right) + \ldots\right]$$

In order to simplify the expression on the right, we find out the number of samples that contain any specified value X_i. The number of such samples is $\binom{N-1}{n-1}$, that is, the number of ways in which the $(n-1)$ other units in the sample are to be selected from the remaining $(N-1)$ units.

Next, we determine the co-efficient of the value X_i by collecting all the terms in the expression containing X_i. Thus the co-efficient of X_i is

$$\frac{\binom{N-1}{n-1}}{k}\cdot\frac{1}{n}=\frac{\binom{N-1}{n-1}}{\binom{N}{n}}\frac{1}{n}=\frac{(N-1)!\,(N-n)!\,n!}{(n-1)!\,(N-n)!\,N!}\cdot\frac{1}{n}=\frac{1}{N}$$

Hence $\mu_{\bar{x}}=\frac{X_1}{N}+\frac{X_2}{N}+\ldots+\frac{X_i}{N}+\ldots+\frac{X_N}{N}$

$$=\frac{X_1+X_2+\ldots+X_i+\ldots+X_N}{N}$$

$= \mu$, mean of the population.

Sampling with replacement. Let $X_1, X_2, \ldots, X_n$ be the observations of a simple random sample of size n from a population having N observations. Then a specified X_i taken from the population, could be any one of the N values with an equal probability of $\frac{1}{N}$ as all the values are equally likely. Thus X_i is a random variable and therefore

$$E(X_i)=\frac{1}{N}\sum_{k=1}^{N}X_k=\mu.$$

For repeated sampling, the mean of a sample $\bar{X}=\frac{1}{n}\sum X_i$ varies from sample to sample, therefore

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

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

$$=\frac{1}{n}\left[\mu+\mu+\ldots+\mu\right]$$

$=\frac{1}{n}[n\mu]=\mu$, the population mean.

(ii) . The standard deviation of the sampling distribution of the mean is given by

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}, \quad (\sigma = \text{population } s.d.)$$

when sampling is performed without replacement from a finite population of size N, *or*

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

when sampling is done with replacement from a finite population or sampling from an infinite population.

Proof. The variance of $\bar{X}$, denoted by $\sigma_{\bar{x}}^2$ or Var($\bar{X}$) is defined as

$$\text{Var}(\bar{X}) = E\,[\bar{X} - E(\bar{X})]^2 = E\,[\bar{X} - \mu]^2$$

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

$$= \frac{1}{n^2} \cdot E\left[\sum_{i=1}^{n}(X_i - \mu)^2 + \sum_{i\neq j}^{n}(X_i - \mu)(X_j - \mu)\right]$$

$$= \frac{1}{n^2}E\left[\sum_{i=1}^{n}(X_i - \mu)^2\right] + \frac{1}{n^2}E\left[\sum_{i\neq j}^{n}(X_i - \mu)(X_j - \mu)\right]$$

The simplification depends on whether the sampling is performed without replacement from a finite population of size N or sampling is done with replacement. The two cases are treated separately.

First case: *Sampling without replacement.*

Since the probability of obtaining $(X_i - \mu)^2$ on the *ith* draw is equal to the probability of obtaining X_i on the *ith* draw which is $\frac{1}{N}$, therefore the expected value of $(X_i - \mu)^2$ becomes σ^2, *i.e.*

$$E(X_i - \mu)^2 = \sum_{i=1}^{N}\frac{1}{N}(X_i - \mu)^2 = \sigma^2.$$

Again, since the sampling is without replacement, the probability of selecting $(X_i - \mu)(X_j - \mu)$ on the *ith* and *jth* draw is $\frac{1}{N} \cdot \frac{1}{N-1}$, because they are not independent on account of the reduction in size from N to $N-1$. Thus

$$E(X_i-\mu)(X_j-\mu) = \frac{1}{N}\cdot\frac{1}{N-1}\sum_{i\neq j}^{N}(X_i-\mu)(X_j-\mu)$$

$$= \frac{1}{N(N-1)}\left\{\left[\sum_{i=1}^{N}(X_i-\mu)\right]^2 - \sum_{i=1}^{N}(X_i-\mu)^2\right\}$$

$$= \frac{1}{N(N-1)}\left\{0 - \sum_{i=1}^{N}(X_i-\mu)^2\right\}$$

$$= \frac{-\sigma^2}{N-1}$$

Substituting these values, we get

$$\text{Var}(\bar{X}) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 + \frac{1}{n^2}\sum_{i\neq j}^{n}\frac{-\sigma^2}{N-1}$$

$$= \frac{1}{n^2}(n\sigma^2) - \frac{1}{n^2}\cdot\frac{1}{N-1}\sum_{i\neq j}^{n}\sigma^2$$

$$= \frac{\sigma^2}{n} - \frac{1}{n^2}\frac{1}{N-1}\cdot n(n-1)\sigma^2$$

$$= \frac{\sigma^2}{n}\cdot\frac{N-n}{N-1}.$$

$$\text{Hence } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\cdot\sqrt{\frac{N-n}{N-1}}.$$

The factor $\frac{N-n}{N-1}$ is usually called the *finite population correction* (*fpc*) or *finite correction factor* (*fcf*) for the variance because in sampling from finite population, the variance of the mean is reduced by this amount. It is important to note that in sampling without replacement from a finite population of size N, *fpc* is dropped from the formula whenever n, the sample size, is less than 5% of N; and *fpc* is used when n is 5% or greater than 5% of N.

Second case : ***Sampling with replacement***

When sampling is done with replacement or sampling from an infinite population, the X_i and X_j are statistically independent. Therefore

$E(X_i - \mu)(X_j - \mu) = 0$. Hence we get

$$\text{Var}(\bar{X}) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n},$$

or $$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\text{population standard deviation}}{\text{square root of sample size}}.$$

It is to be noted that the standard error of the mean is always less than the standard deviation of the population. This means that the sampling distribution of the mean has less variability than the population from which the samples were taken. If the value of σ is not known and if the sample size is *large* (as a rule of thumb adopted by many authors, a sample containing 30 or more observations constitutes a *large* or *sufficiently large* sample), it is replaced by s, the standard deviation of the sample. The *S.E.* of the mean then becomes

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}.$$

(iii) **Shape of the distribution.** (a) If the population sampled is normally distributed, then the sampling distribution of the mean $\bar{X}$, will also be normal regardless of sample size.

To prove this, we proceed as follows:

By definition, the moment generating function of $\bar{X}$ is

$$M_{\bar{X}}(t) = E(e^{t\bar{X}}) = E(e^{\Sigma tX_i/n})$$

$$= E\left[\prod_{i=1}^{n} e^{tX_i/n}\right] = \prod_{i=1}^{n} E(e^{tX_i/n}).$$

But $$E(e^{tX_i/n}) = M_X\left(\frac{t}{n}\right).$$

If X is $N(\mu, \sigma^2)$, then

$$M_X(t) = e^{\mu t + \sigma^2 t^2/2}, \text{ and}$$

$$M_X\left(\frac{t}{n}\right) = e^{\mu t/n + \frac{1}{2}\sigma^2(t/n)^2}$$

Since $X_1, X_2, \ldots, X_n$ is a random sample, therefore

$$M_{\bar{X}}(t) = E\left(e^{t\bar{X}}\right) = \left(E\left(e^{tX_i/n}\right)\right)^n$$

$$= \left[M_X\left(\frac{t}{n}\right)\right]^n = e^{n\left(\frac{\mu t}{n} + \frac{1}{2}\sigma^2 t^2/n^2\right)}$$

$$= e^{\mu t + \sigma^2 t^2/2n}$$

But this is the *m.g.f.* of a normal distribution with mean = μ and variance = $\frac{\sigma^2}{n}$. Thus $\bar{X}$ is normally distributed variable with mean μ and variance σ^2/n where μ and σ^2 are the mean and variance of the population.

(b) If the population sampled is non-normal, then for sufficiently large sample size, the sampling distribution of $\bar{X}$ will approximate the normal distribution.

This is a special case of the most important statistical theorem, known as the *Central Limit Theorem*, which is stated and proved in the next section.

We know that the standardized form of a random variable is obtained by subtracting its mean from it and dividing the difference by its standard deviation, that is

$$Z = \frac{\text{value of random variable} - \text{mean of random variable}}{\text{standard deviation of random variable}}$$

We have proved above that the sample mean $\bar{X}$ is normally distributed random variable with mean equal to population mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$. The standard normal variable then becomes

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

If sampling is without replacement and sample size n is 5% or greater than 5 per cent of the population size N, then Z values are obtained by the formula

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}}$$

The sampling distribution of $\bar{X}$ thus offers solutions to probability questions about the values of the sample means.

Example 14.7. Assume that a population consists of 7 similar containers having the following weights (kilograms):

9.8, 10.2, 10.4, 9.8, 10.0, 10.2 and 9.6.

(a) Find the mean μ and the standard deviation σ of the given population.

(b) Draw random samples of 2 containers without replacement and calculate the mean weight $\bar{X}$ of each sample.

(c) Form a frequency distribution of $\bar{X}$ and a sampling distribution of $\bar{X}$.

(d) Find the mean and the standard deviation of the sampling distribution of $\bar{X}$.

(a) The population mean μ and standard deviation σ are

$$\mu = \frac{\Sigma X}{N} = \frac{9.8 + 10.2 + \dots + 9.6}{7} = \frac{70.0}{7} = 10.0 \text{ kg; and}$$

$$\sigma = \sqrt{\frac{\Sigma(X-\mu)^2}{N}} = \sqrt{\frac{(9.8-10)^2+(10.2-10)^2+\dots+(9.6-10)^2}{7}}$$

$$= \sqrt{\frac{0.48}{7}} = \sqrt{0.0686} = 0.262 \text{ kg.}$$

(b) Let the containers be identified as A, B, C, D, E, F and G. Now the number of possible random samples of $n = 2$ containers without replacement is $\binom{7}{2} = 21$. The 21 possible random samples with the values of their mean weights are given on page 32:

(c) The frequency distribution of $\bar{X}$ and the sampling distribution of the mean $\bar{X}$, which is just the relative frequency distribution of $\bar{X}$, are obtained below:

(i) Frequency Distribution of $\bar{X}$

Sample Mean $\bar{x}$	Tally	f
9.7	II	2
9.8	II	2
9.9	IIII	4
10.0	~~IIII~~	5
10.1	IIII	4
10.2	II	2
10.3	II	2
Σ	--	21

(ii) Sampling Distribution of $\bar{X}$

Sample Mean $\bar{x}$	Probability $f(\bar{x})$
9.7	2/21
9.8	2/21
9.9	4/21
10.0	5/21
10.1	4/21
10.2	2/21
10.3	2/21
Σ	1

Sample No.	Sample Combination	Weights in X_1	Samples X_2	Sample Mean weight ($\bar{X}$)
1	*A, B*	9.8,	10.2	10.0
2	*A, C*	9.8,	10.4	10.1
3	*A, D*	9.8,	9.8	9.8
4	*A, E*	9.8,	10.0	9.9
5	*A, F*	9.8,	10.2	10.0
6	*A, G*	9.8,	9.6	9.7
7	*B, C*	10.2,	10.4	10.3
8	*B, D*	10.2,	9.8	10.0
9	*B, E*	10.2,	10.0	10.1
10	*B, F*	10.2,	10.2	10.2
11	*B, G*	10.2,	9.6	9.9
12	*C, D*	10.4,	9.8	10.1
13	*C, E*	10.4,	10.0	10.2
14	*C, F*	10.4,	10.2	10.3
15	*C, G*	10.4,	9.6	10.0
16	*D, E*	9.8,	10.0	9.9
17	*D, F*	9.8,	10.2	10.0
18	*D, G*	9.8,	9.6	9.7
19	*E, F*	10.0,	10.2	10.1
20	*E, G*	10.0,	9.6	9.8
21	*F, G*	10.2,	9.6	9.9

(d) The mean and standard deviation of sampling distribution of $\bar{X}$, are computed below:

Sample Mean $\bar{x}$	Probability $f(\bar{x})$	$\bar{x} f(\bar{x})$	$\bar{x}-\mu_{\bar{x}}$	$(\bar{x}-\mu_{\bar{x}})^2$	$(\bar{x}-\mu_{\bar{x}})^2 f(\bar{x})$
9.7	2/21	19.4/21	−0.3	0.09	0.18/21
9.8	2/21	19.6/21	−0.2	0.04	0.08/21
9.9	4/21	39.6/21	−0.1	0.01	0.04/21
10.0	5/21	50.0/21	0	0	0
10.1	4/21	40.4/21	0.1	0.01	0.04/21
10.2	2/21	20.4/21	0.2	0.04	0.08/21
10.3	2/21	20.6/21	0.3	0.09	0.18/21
Σ	1	10.0	--	--	0.6/21

$$\mu_{\bar{x}} = \Sigma \bar{x} f(\bar{x}) = 10.0 \; kg, \text{ and}$$

$$\sigma_{\bar{x}} = \sqrt{\Sigma (\bar{x} - \mu_{\bar{x}})^2 f(\bar{x})} = \sqrt{\frac{0.6}{21}} = \sqrt{0.0286} = 0.17 \; kg,$$

which is a smaller value indicating that the sampling distribution of the mean is more concentrated about the population mean.

Example 14.8. A sample of size $n=3$ is to be randomly selected without replacement from a population that has $N=5$ items whose values are 0, 3, 6, 9 and 12.

(a) Find the sampling distribution of the sample mean, $\bar{X}$.

(b) Calculate the mean and the standard deviation of $\bar{X}$, and verify that

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}.$$

Let the items be designated by the letters *A*, *B*, *C*, *D* and *E*.

(a) The number of samples of size $n=3$ that could be drawn without replacement from a population of size $N=5$ is

$$\binom{5}{3} = \frac{5!}{2!\,3!} = 10.$$

The 10 possible samples and their means are given below:

Sample No.	Sample Combinations	Sample Values	Sample Mean ($\bar{X}$)
1	*A, B, C*	0, 3, 6	3
2	*A, B, D*	0, 3, 9	4
3	*A, B, E*	0, 3, 12	5
4	*A, C, D*	0, 6, 9	5
5	*A, C, E*	0, 6, 12	6
6	*A, D, E*	0, 9, 12	7
7	*B, C, D*	3, 6, 9	6
8	*B, C, E*	3, 6, 12	7
9	*B, D, E*	3, 9, 12	8
10	*C, D, E*	6, 9, 12	9

The sampling distribution is obtained by listing all possible means and their probabilities (relative frequencies) as below:

Sampling Distribution of $\bar{X}$

Sample Mean $\bar{X}$	Number of sample means (f)	Probability $f(\bar{x})$
3	1	1/10
4	1	1/10
5	2	2/10
6	2	2/10
7	2	2/10
8	1	1/10
9	1	1/10
Σ	10	1

(b) Next, we calculate the mean and the standard deviation (the *standard error*) of the sampling distribution of the mean as follows:

Calculation of Mean and S.D. of Sampling Distribution of $\bar{X}$.

Sample Mean $\bar{x}$	Probability $f(\bar{x})$	$\bar{x}\,f(\bar{x})$	$\bar{x}^2\,f(\bar{x})$
3	1/10	3/10	9/10
4	1/10	4/10	16/10
5	2/10	10/10	50/10
6	2/10	12/10	72/10
7	2/10	14/10	98/10
8	1/10	8/10	64/10
9	1/10	9/10	81/10
Σ	1	60/10	390/10

Now $\mu_{\bar{x}} = \sum \bar{x}\,f(\bar{x}) = \dfrac{60}{10} = 6$, and

$$\sigma_{\bar{x}} = \sqrt{[\Sigma\bar{x}^2 f(\bar{x})] - [\Sigma\,\bar{x}f(\bar{x})]^2}$$

$$= \sqrt{\frac{390}{10} - \left(\frac{60}{10}\right)^2} = \sqrt{39 - 36} = \sqrt{3} = 1.732$$

In order to verify the given result, we first calculate the mean μ and the variance σ^2 of the given population. Thus

$$\mu = \frac{1}{5}[0 + 3 + 6 + 9 + 12] = \frac{1}{5}[30] = 6, \text{ and}$$

$$\sigma^2 = \frac{1}{5}[(0-6)^2 + (3-6)^2 + (6-6)^2 + (9-6)^2 + (12-6)^2] = 18$$

Verification: Now $\dfrac{\sigma^2}{n} \cdot \dfrac{N-n}{N-1} = \dfrac{18}{3} \cdot \dfrac{5-3}{5-1} = \dfrac{18\times2}{3\times4} = 3 = \sigma_{\bar{x}}^2$

Hence the result.

Example 14.9. Suppose that a random variable X has the following population distribution:

x	3	6	9
$f(x)$	1/3	1/3	1/3

If a sample of three numbers is taken with replacement, obtain the sampling distribution of the sample mean and verify that $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$.

The population distribution of the r.v. X may be written as

$$f(x) = P(X=x) = \begin{cases} 1/3 & \text{for } x = 3 \\ 1/3 & \text{for } x = 6 \\ 1/3 & \text{for } x = 9 \end{cases}$$

implying that the members of the population have the numerical values of 3, 6 and 9.

The number of possible samples of size $n=3$, which could be selected with replacement from this population is $3^3=27$. The 27 random samples with their means are given below:

Sample No.	Sample Values X_1	X_2	X_3	Sample Mean $\bar{X}$
1	3	3	3	3
2	3	3	6	4
3	3	3	9	5
4	3	6	3	4
5	3	6	6	5
6	3	6	9	6
7	3	9	3	5
8	3	9	6	6
9	3	9	9	7
10	6	3	3	4
11	6	3	6	5
12	6	3	9	6
13	6	6	3	5
14	6	6	6	6
15	6	6	9	7
16	6	9	3	6
17	6	9	6	7
18	6	9	9	8
19	9	3	3	5
20	9	3	6	6
21	9	3	9	7
22	9	6	3	6
23	9	6	6	7
24	9	6	9	8
25	9	9	3	7
26	9	9	6	8
27	9	9	9	9

The sampling distribution of the sample mean $\bar{X}$ is obtained below, together with two columns needed for the calculation of the S.E. of the mean:

Sampling Distribution of $\bar{X}$ and Calculation of $S.E. (\bar{X})$

$\bar{X}$	No. of sample means	Probability $f(\bar{x})$	$\bar{x} f(\bar{x})$	$\bar{x}^2 f(\bar{x})$
3	1	1/27	3/27	9/27
4	3	3/27	12/27	48/27
5	6	6/27	30/27	150/27
6	7	7/27	42/27	252/27
7	6	6/27	42/27	294/27
8	3	3/27	24/27	192/27
9	1	1/27	9/27	81/27
Σ	27	1	162/27	1026/27

Now, $\mu_{\bar{x}} = \sum \bar{x} f(\bar{x}) = \dfrac{162}{27} = 6$

$$\sigma_{\bar{x}}^2 = \sum \bar{x}^2 f(\bar{x}) - [\sum \bar{x} f(\bar{x})]^2$$

$$= \frac{1026}{27} - (6)^2 = 38 - 36 = 2.$$

And, $\mu = \sum x f(x) = 3 \times \frac{1}{3} + 6 \times \frac{1}{3} + 9 \times \frac{1}{3} = 6,$

$$\sigma^2 = \sum x^2 f(x) - \mu^2 = \left[9 \times \frac{1}{3} + 36 \times \frac{1}{3} + 81 \times \frac{1}{3}\right] - (6)^2$$

$$= (3 + 12 + 27) - 36 = 6.$$

Verification: $\dfrac{\sigma^2}{n} = \dfrac{6}{3} = 2 = \sigma_{\bar{x}}^2$

Hence, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$.

Example 14.10. The weights of 1500 ball bearings are normally distributed with a mean of 22.40 ounces and a standard deviation of 0.048 ounces. If 300 random samples of size 36 are drawn from this population, (a) determine the expected mean and standard deviation of the sampling distribution of mean if sampling is done (i) with replacement, (ii) without replacement; (P.U. B.A./B.Sc., 1971)

(b) How many of the random samples would have their means between 22.39 and 22.42 oz?

(a) There would be $(1500)^{36}$ and $\binom{1500}{36}$ possible samples of size 36, that could be obtained theoretically from a population of weights of 1500 ball bearings with and without replacement respectively. Obviously the number of theoretically possible samples is much larger than 300. Therefore the sampling distribution of the mean will not be a true sampling distribution. (Such a distribution is called *experimental* sampling distribution). But 300 being a large number, there should be a close agreement between the experimental sampling distribution of 300 sample means and the true sampling distribution of mean. Hence the expected mean and standard deviation are found to be as:

(i) *Sampling with replacement:*

$$\mu_{\bar{x}} = \mu = 22.40 \text{ oz.},$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.048}{\sqrt{36}} = 0.008 \text{ oz.}$$

(ii) *Sampling without replacement:*

$$\mu_{\bar{x}} = \mu = 22.40 \text{ oz.},$$

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

Since sample size $n=36$ is less than 5% of the population size $N=1500$, therefore according to the generally accepted rule for the use of *fpc*, the factor $\sqrt{\frac{N-n}{N-1}}$ is dropped. Thus

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.048}{\sqrt{36}} = 0.008 \text{ oz.}$$

(b) The sampling distribution of the mean $\bar{X}$ is normal because the population sampled is normally distributed. Thus

$$Z = \frac{\bar{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - 22.40}{0.008}$$ is a standard normal variable.

To find the expected number of samples that would have their means between 22.39 and 22.42 *oz*, we will transform these values to the corresponding *z*-values. Thus

$$\text{at } \bar{x} = 22.39, \text{ we find } z_1 = \frac{22.39 - 22.40}{0.008} = -1.25, \text{ and}$$

$$\text{at } \bar{x} = 22.42, \text{ we find } z_2 = \frac{22.42 - 22.40}{0.008} = 2.50.$$

Using the **Table** of areas under the normal curve, we find

$$P(22.39 \le \bar{X} \le 22.42) = P(-1.25 \le Z \le 2.50)$$
$$= P(-1.25 \le Z \le 0) + P(0 \le Z \le 2.50)$$
$$= 0.3944 + 0.4938 = 0.8882.$$

Hence the expected number of samples = (300) (0.8882) = 267.

Example 14.11. A construction company has 310 employees who have an average annual salary of Rs. 24,000. The standard deviation of annual salaries is Rs. 5,000. In a random sample of 100 employees, what is the probability that the average salary will exceed Rs. 24,500?

The sample size ($n = 100$) is large enough to assume that the sampling distribution of $\bar{X}$ is approximately normally distributed with mean

$$\mu_{\bar{x}} = \mu = \text{Rs. } 24{,}000.$$

and standard deviation

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}} = \frac{5000}{\sqrt{100}} \sqrt{\frac{310-100}{310-1}}$$
$$= \text{Rs. } 412.20,$$

where we have used *fpc*, because the sample size $n = 100$ is greater than 5 per cent of the population size $N = 310$.

$$\text{Equivalently, } Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X} - 24000}{412.20} \text{ is approximately } N(0, 1).$$

We are required to evaluate $P(\bar{X} > 24{,}500)$.

$$\text{At } \bar{x} = 24{,}500, \text{ we find that } z = \frac{24500 - 24000}{412.20} = 1.21$$

Hence using *Table* of areas under normal curve, we get

$$P(\bar{X} > 24{,}500) = P(Z > 1.21)$$
$$= 0.5 - P(0 \le Z \le 1.21)$$
$$= 0.5 - 0.3869 = 0.1131.$$

Example 14.12. Calculate the standard error of the mean from the following data collected in one of the many random sample inquiries to find average earning of a particular class.

Earning (Rs.)	1–10	11–20	21–30	31–40	41–50	51–60	61–70	71–80
Number of persons	20	98	150	218	200	164	110	40

Since the population standard deviation σ is not known and the sample size ($n=1000$) is large enough to replace it with the sample standard deviation s, we therefore first calculate the sample standard deviation as below:

Earning (Rs.)	f	u	fu	fu^2
1–10	20	–3	–60	180
11–20	98	–2	–196	392
21–30	150	–1	–150	150
31–40	218	0	0	0
41–50	200	1	200	200
51–60	164	2	328	656
61–70	110	3	330	990
71–80	40	4	160	640
Σ	1000	--	612	3208

The sample mean and the sample standard deviation are;

$$\bar{x} = a + \frac{\Sigma fu}{n} \times h = 35.5 + \frac{612}{1000} \times 10 = \text{Rs. } 41.62, \text{ and}$$

$$s = h\sqrt{\frac{\Sigma fu^2}{n} - \left(\frac{\Sigma fu}{n}\right)^2}$$

$$= 10\sqrt{\frac{3208}{1000} - \left(\frac{612}{1000}\right)^2} = 10\sqrt{2.833456} = \text{Rs. } 16.83.$$

Hence the standard error of the sample mean is

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{16.83}{\sqrt{1000}} = \frac{16.83}{31.62} = \text{Rs. } 0.53.$$