Note: Sample variance (s^2)

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

Note: Here, in sample variance, we have divided it by $(n-1)$ in place of n because if we divide it by 'n' then sample variance will not be unbiased for population variance and we will loose very important characteristic in statistics. Since, it is little bit deep concept so just skip these things and remember the formula of sample-variance.

population

Total set of observations or collection of some measurements is known as population.

Sample: A subset of population is known as sample, and the process of selecting sample from population is known as sampling.

Q. Let $x_1, x_2, \cdots, x_n$ are iid and follow the $N(\mu, \sigma^2)$ where μ and σ^2 are characteristic of population then find the sampling distribution of $\bar{x}$.

Sol^n

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

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

Note: We know that if X_i; $i = 1, 2, \cdots, n$ follows the normal distribution then any linear combination of these X_i also follows the normal dist^n (We prove it Later).

Now, we have to find the distribution of $\bar{x}$

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

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

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

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

$$= \frac{1}{n}\underbrace{(\mu + \mu + \cdots + \mu)}_{(n\text{-times})} = \frac{\cancel{n}\mu}{\cancel{n}} = \mu.$$

Now, variance of $\bar{x}$,

$$V(\bar{X}) = V\left(\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right) \quad \left(\begin{array}{c} V(aX) \\ \parallel \\ a^2 V(X) \end{array}\right)$$

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

$$= \frac{1}{n^2}(\sigma^2 + \sigma^2 + \cdots + \sigma^2)_{(n\text{-times})}$$

$$V(\bar{X}) = \frac{\cancel{n}\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$\therefore$ Thus, $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

Result: If $X_i \sim N(\mu, \sigma^2)$ and X_i are iid then any linear combination also follow the same distribution.

proof Consider $Z = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$

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

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

$$= a_1 \mu + a_2 \mu + \cdots + a_n \mu$$

$$= (a_1 + a_2 + \cdots + a_n) \mu$$

$$V(Z) = V(a_1 X_1 + a_2 X_2 + \cdots + a_n X_n)$$

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

$$= (a_1^2 + a_2^2 + \cdots + a_n^2) \sigma^2$$

Thus; $Z \sim N\left((a_1 + a_2 + \cdots + a_n)\mu, (a_1^2 + a_2^2 + \cdots + a_n^2)\sigma^2 \right)$.

Note Each X_i are independent and have the same distribution (iid), then $Cov(x_i, x_j) = 0 \ \forall \ i \neq j$.