7. The mean and standard deviation of a set of n_1 observations are $\bar{x}_1$ and s_1, respectively while the mean and standard deviation of another set of n_2 observations are $\bar{x}_2$ and s_2, respectively. Show that the standard deviation of the combined set of $(n_1 + n_2)$ observations is given by

$$\therefore \qquad \text{SD} = \sqrt{\frac{n_1(s_1)^2 + n_2(s_2)^2}{n_1 + n_2} + \frac{n_1 n_2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2}}$$

Sol. We have two sets of observations,

$$x_i, i = 1, 2, 3 \ldots, n_1 \text{ and } y_j, j = 1, 2, 3, \ldots, n_2$$

$$\therefore \qquad \bar{x}_1 = \frac{1}{n_i}\sum_{i=1}^{n_1} x_i \text{ and } \bar{x}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} y_j$$

$$\Rightarrow \qquad \sigma_1^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2 \text{ and } \sigma_2^2 = \frac{1}{n_2}\sum_{j=1}^{n_2}(y_i - \bar{x}_2)^2$$

Now, mean $\bar{x}$ of the given series is given by

$$\bar{x} = \frac{1}{n_1 + n_2}\left[\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} y_j\right] = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

The variance σ^2 of the combined series is given by

$$\sigma^2 = \frac{1}{n_1 + n_2}\left[\sum_{i=1}^{n_1}(x_i - \bar{x})^2 + \sum_{j=1}^{n_2}(y_j - \bar{x})^2\right]$$

Now, $\sum_{i=1}^{n_1}(x_i - \bar{x})^2 = \sum_{i=1}^{n_1}(x_i - \bar{x}_j + \bar{x}_j - \bar{x})^2$

But $\sum_{i=1}^{n_1}(x_i - \bar{x}_i) = 0$ [Algebraic sum of the deviation of values of first series from their mean is zero]