In the formula, we add the factor $\frac{m(m^2-1)}{12}$ to Σd^2, where m is the number of times an item is repeated. This correction factor is to be added for each repeated value in both the X-series and Y-series.

Example 10·18. *Obtain the rank correlation coefficient for the following data:*

X	:	68	64	75	50	64	80	75	40	55	64
Y	:	62	58	68	45	81	60	68	48	50	70

Solution.

CALCULATIONS FOR RANK CORRELATION

X	Y	Rank X (x)	Rank Y (y)	$d = x - y$	d^2
68	62	4	5	−1	1
64	58	6	7	−1	1
75	68	2·5	3·5	−1	1
50	45	9	10	−1	1
64	81	6	1	5	25
80	60	1	6	−5	25
75	68	2·5	3·5	−1	1
40	48	10	9	1	1
55	50	8	8	0	0
64	70	6	2	4	16
				$\Sigma d = 0$	$\Sigma d^2 = 72$

In the X-series we see that the value 75 occurs 2 times. The common rank given to these values is 2·5 which is the average of 2 and 3, the ranks which these values would have taken if they were different. The next value 68, then gets the next rank which is 4. Again we see that value 64 occurs thrice. The common rank given to it is 6 which is the average of 5, 6 and 7. Similarly in

the Y-series, the value 68 occurs twice and its common rank is 3·5 which is the average of 3 and 4. As a result of these common rankings, the formula for 'ρ' has to be corrected. To Σd^2 we add $\frac{m(m^2-1)}{12}$ for each value repeated, where m is the number of times a value occurs. In the X-series the correction is to be applied twice, once for the value 75 which occurs twice ($m = 2$) and then for the value 64 which occurs thrice ($m = 3$). The total correction for the X-series is

$$\frac{2(4-1)}{12} + \frac{3(9-1)}{12} = \frac{5}{2}$$

Similarly, this correction for the Y-series is $\frac{2(4-1)}{12} = \frac{1}{2}$, as the value 68 occurs twice.

$$\text{Thus} \quad \rho = 1 - \frac{6\left[\Sigma d^2 + \frac{5}{2} + \frac{1}{2}\right]}{n(n^2-1)} = 1 - \frac{6(72+3)}{10 \times 99} = 0{\cdot}545$$

10·6·3. Limits for the Rank Correlation Coefficient. Spearman's rank correlation coefficient is given by

$$\rho = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

'ρ' is maximum, if $\sum_{i=1}^{n} d_i^2$ is minimum, *i.e.*, if each of the deviations d_i is minimum. But the minimum value of d_i is zero in the particular case $x_i = y_i$, *i.e.*, if the ranks of the ith individual in the two characteristic are equal. Hence the maximum value of ρ is + 1, *i.e.*, $\rho \leq 1$.

'ρ' is minimum , if $\sum_{i=1}^{n} d_i^2$ is maximum, *i.e.*, if each of the deviations d_i is maximum which is so if the ranks of the n individuals in the two