

где n_{ij} - частоты пар (x_i, y_j) и $n_i = \sum_{j=1}^m n_{ij}$; m - число интервалов по переменной Y . Получим

$$\bar{y}_1 = \frac{20 \cdot 17 + 30 \cdot 4}{21} = \frac{460}{21} = 21,905;$$

$$\bar{y}_2 = \frac{20 \cdot 3 + 30 \cdot 18 + 40 \cdot 3}{24} = \frac{720}{24} = 30;$$

$$\bar{y}_3 = \frac{30 \cdot 2 + 40 \cdot 15 + 50 \cdot 5}{22} = \frac{910}{22} = 41,364;$$

$$\bar{y}_4 = \frac{40 \cdot 3 + 50 \cdot 13 + 60 \cdot 7}{23} = \frac{1190}{23} = 51,739;$$

$$\bar{y}_5 = \frac{60 \cdot 6 + 70 \cdot 14}{20} = \frac{1340}{20} = 67.$$

Вычисленные групповые средние \bar{y}_j поместим в последнем столбце корреляционной таблицы и изобразим графически в виде ломанной, называемой эмпирической линией регрессии Y по X .