

相关性是协方差的标准化格式。协方差本身很难做比较。例如：如果我们计算工资（\$）和年龄（岁）的协方差，因为这两个变量有不同的度量，所以我们会得到不能做比较的不同的协方差。

$$\Sigma_{ij} = \text{cov}(X_i, X_j) = \mathbf{E}[(X_i - \mu_i)(X_j - \mu_j)] = \mathbf{E}[X_i X_j] - \mu_i \mu_j$$

为了解决这个问题，我们计算相关性来得到一个介于-1和1之间的值，就可以忽略它们各自不同的度量。

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$