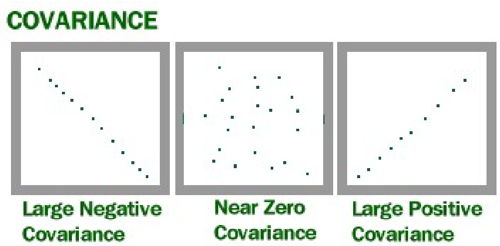
Covariance

Statistics > Covariance. Covariance is a measure of how much two random variables vary together. It's similar to variance, but where variance tells you how a single variable varies, co variance tells you how two variables vary together.



What is the use of covariance?

Covariance is a measure of how changes in one variable are associated with changes in a second variable. Specifically, covariance measures the degree to which two variables are linearly associated. However, it is also often used informally as a general measure of how monotonically related two variables are.

What does covariance between variables mean?

In probability theory and statistics, covariance is a measure of the joint variability of two random variables. ... In the opposite case, when the greater values of one variable mainly correspond to the lesser values of the other, i.e., the variables tend to show opposite behavior, the covariance is negative.

- a =

4.0000	2.0000	0.6000
4.2000	2.1000	0.5900
3.9000	2.0000	0.5800
4.3000	2.1000	0.6200
4.1000	2.2000	0.6300

- The set of 5 observations, measuring 3 variables, can be described by its *mean vector* and *variance-covariance matrix*. The three variables, from left to right are length, width, and height of a certain object, for example. Each row vector is another observation of the three variables (or components).

- *Definition of mean vector and covariance matrix*
- The mean vector consists of the means of each variable and the covariance matrix consists of the variances of the variables along the main diagonal and the covariances between each pair of variables in the other matrix positions.

- S is calculated by

$$S = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})',$$

where n=5 for this example, sometimes for practical reasons instead of n-1 we use n in the first denominator.

Covariance
For two random variable vectors A and B, the covariance is defined as

$$\mathrm{cov}(A,B) = \frac{1}{N-1}\sum_{i=1}^{N}(A_i - \mu_A)^*(B_i - \mu_B)$$

where µA is the mean of A, µB is the mean of B, and * denotes the complex conjugate.

- mean(a)

ans =

4.1000 2.0800 0.6040

- >> cov(a)

ans =

0.0250	0.0075	0.0017
0.0075	0.0070	0.0014
0.0017	0.0014	0.0004

Thus, 0.025 is the variance of the length variable, 0.0075 is the covariance between the length and the width variables, 0.00175 is the covariance between the length and the height variables, 0.007 is the variance of the width variable, 0.00135 is the covariance between the width and height variables and 0.00043 is the variance of the height variable.