COVARIANCE

1. Introduction

- Covariance is number which tells us how two variables vary together, if and how they are related, how do they co-vary.
- Covariance tells us that if we increase first variable will the other one get bigger, smaller or will it stay the same.
- Covariance value is calculated like this [2]:

$$\mathrm{cov}(x, y) = E[(x - \bar{x})(y - \bar{y})] \tag{1.1}$$

$$\mathrm{cov}(x, y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}) \tag{1.2}$$

where:
- cov(x,y) – covariance between variables x and y,
- $\bar{x}$, $\bar{y}$ – mean values [3] of variables x and y respectively,
- E(z) – mean values [3] of variable z,
- N – number of measurments taken for variables x and y,
- x_i, y_i – value of variables x and y respectively taken at i-th measurement.

- Following figure [1] shows few examples of different covariances between variables x and y where ach dot in represents specific moment at which we have measured values of variables x and y:

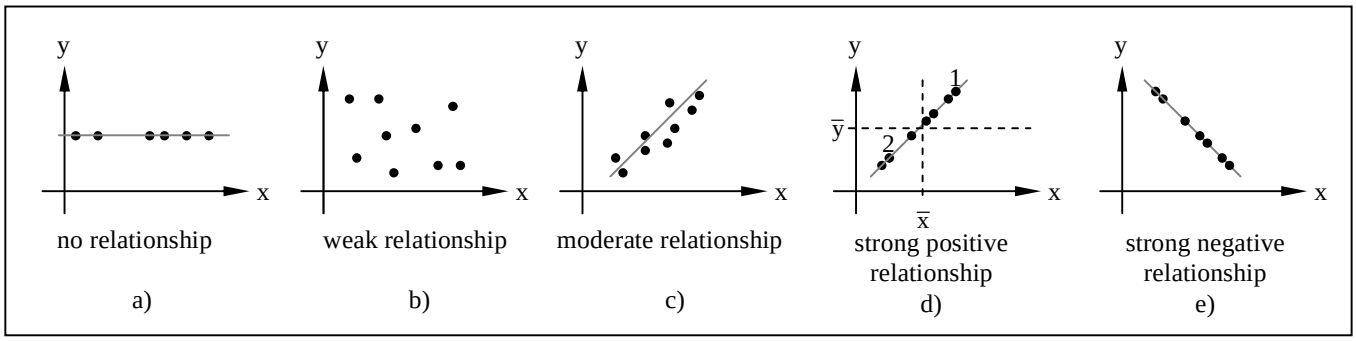

Figure 1.1 Examples of different covariances between varaibles x and y.

- When there is no relationship between variables, like in example on figure 1.1 a), it means that changing value of variable x will have no effect on value of variable y, therefore variable y stays the same.
  In other words, if you know value of x there is no way to calculate value of y.
- When there is weak or moderate relationship between variables, like in example on figures 1.1 b) and c), it means that changing value of variable x will have week effect on variable y, because variable y depends more on some other variables that were changed between measurments then on values of x.
  In other words, if you know value of x you can presume that value of y will be inside some interval.
- When there is strong relationship between variables, like in examples on figures 1.1 d) and e), it means that changing value of variable x will have great effect on variable y, since y mostly depends on x.
  In other words, if you know value of x you can then you can exactly calculate the value of y.

- Positive covariance means that if you increase one variable, the other variable y will also get increased, and other way around
- Negative covariance means that if you increase one variable, the other variable y will get decreased, and other way around.
- Zero covariance means that variables are unrelated, which meand that changing one variable has no effect on the other.

- Calculating (1.2) for dot 1 on figure 1.1 d), both values under round brackets will be positive, so the result will also be positive.
  Doing the same for dot 2, both values under round brackets will be negative, so the result is again positive.
  Resulting covariance will be therefore positive, indicating that variables grow and decay together as it is the case.

2. References

[1] http://www.jsu.edu/depart/psychology/sebac/fac-sch/rm/Ch3-2.html

[2] http://www.netmba.com/statistics/covariance/

[3] Mean value.pdf