Covariance and Correlation – Solutions
STAT-UB.0103 – Statistics for Business Control and Regression Models

Portfolio Optimization

1. Suppose there are two stocks, X and Y. The annual returns for these stocks can be modeled as independent random variables. Suppose that the expected returns for the two stocks are both equal to 5%, and the standard deviations of the returns for the two stocks are both equal to 1%. Suppose you invest \$30 in stock X and \$70 in stock X.

   (a) If after one year, the return from stock X is 6.0% and the return from stock Y is 4.8%, what is your gain?

   Solution: Let X denote the return from stock X, and Y denote the return from stock Y. Your gain (in dollars) is equal to

   $$G = 30X + 70Y.$$

   Thus, if $X = .060$ and $Y = .048$, then your gain is

   $$G = 30 \times .060 + 70 \times .048 = 5.16.$$

   (b) Suppose you do not know what the returns will be one year from now, and you take them to be random. What is your expected gain?

   Solution: The random variables X and Y have means and standard deviations

   $$\begin{aligned} \mu_X &= \mu_Y = .05, \\ \sigma_X &= \sigma_Y = .01. \end{aligned}$$

   Thus, your expected gain is

   $$\begin{aligned} \mathrm{E}[G] &= \mathrm{E}[30X + 70Y] \\ &= 30\,\mathrm{E}[X] + 70\,\mathrm{E}[Y] \\ &= 30 \times .05 + 70 \times .05 \\ &= 5. \end{aligned}$$

   (c) Suppose I give you \$100 to invest in stocks X and Y. List some strategies for splitting the money between these two stocks.

   Solution: There are two obvious strategies: A, invest all of the money in stock X; and B, split the money evenly between stock X and stock Y. The gains from each of these strategies are

   $$\begin{aligned} A &= 100X, \\ B &= 50X + 50Y. \end{aligned}$$

(d) What are the expected gains from the strategies you devised in part (c)?

Solution: The expected gains are

$$\begin{aligned} \mathrm{E}[A] &= 100\,\mathrm{E}[X] = 100 \times .05 = 5, \\ \mathrm{E}[B] &= 50\,\mathrm{E}[X] + 50\,\mathrm{E}[Y] = 50 \times .05 + 50 \times .05 = 5. \end{aligned}$$

In fact, one can show that no matter how you allocate the money between these two stocks, the expected returns will be the same.

(e) Is there any difference between these investment strategies? Which one should you choose?

Solution:
The gains from each strategy have different standard deviations (referred to as *volatilities* in the finance community). We can compute the standard deviations by first computing the variances and then taking the square roots:

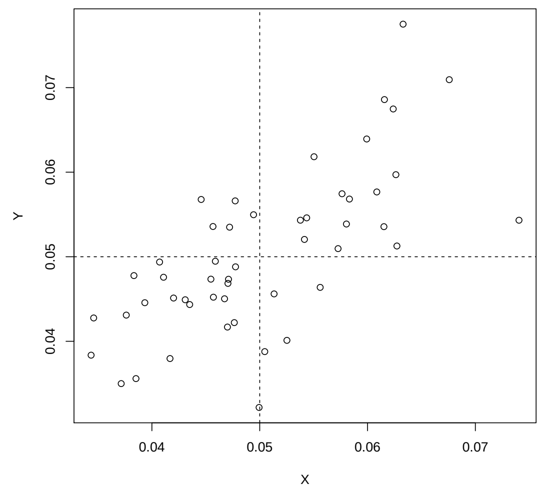
$$\begin{aligned} \mathrm{var}(A) &= \mathrm{var}(100X) \\ &= 100^2\,\mathrm{var}(X) \\ &= 100^2 \times (.01)^2 \\ &= 1. \\ \mathrm{var}(B) &= \mathrm{var}(50X + 50Y) \\ &= \mathrm{var}(50X) + \mathrm{var}(50Y) \\ &= 50^2\,\mathrm{var}(X) + 50^2\,\mathrm{var}(Y) \\ &= 50^2 \times (.01)^2 + 50^2 \times (.01)^2 \\ &= 0.5. \end{aligned}$$

Thus,

$$\begin{aligned} \mathrm{sd}(A) &= \sqrt{\mathrm{var}(A)} = 1, \\ \mathrm{sd}(B) &= \sqrt{\mathrm{var}(B)} \approx .71. \end{aligned}$$

We can see that strategy B has a smaller standard deviation, so the returns from that strategy will tend to be closer to the expected value (i.e. there is less risk).

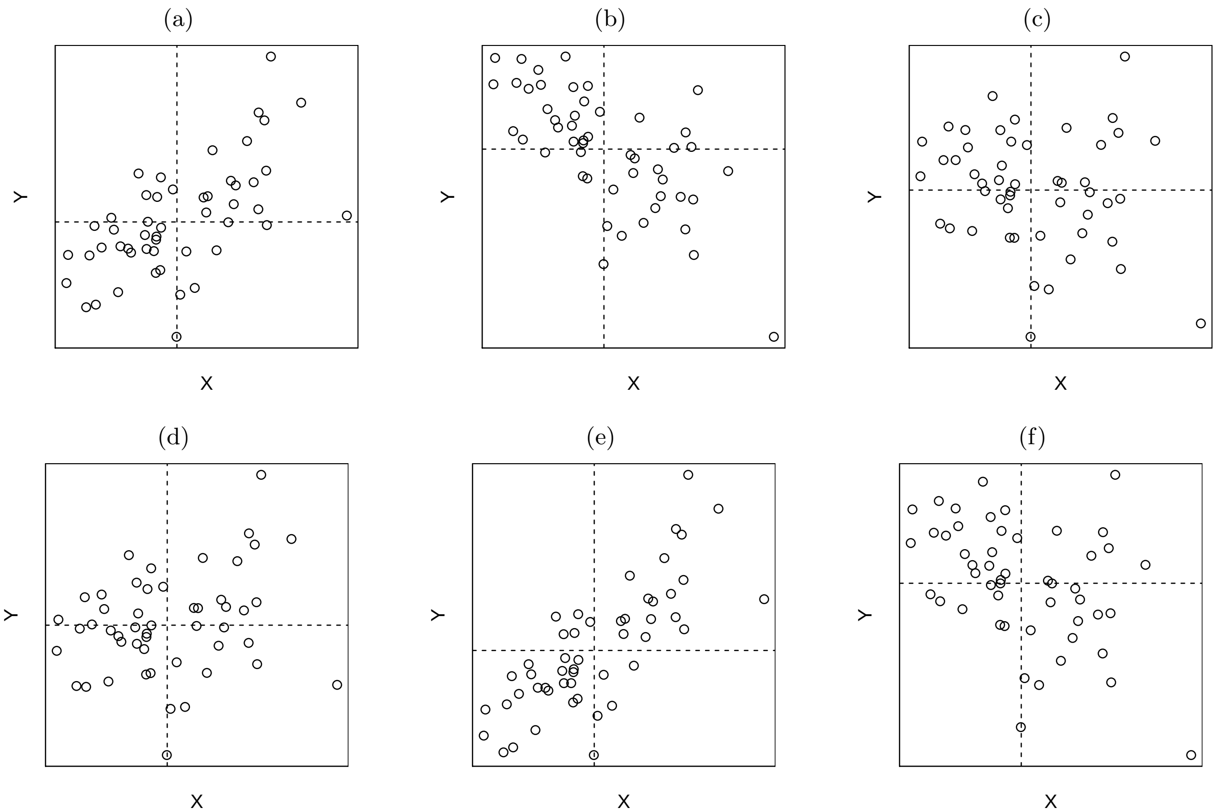
(f) Here is a plot of the annual returns for stocks X and Y for the last fifty years. The means for the two stocks are shown by the dashed lines. Does this plot indicate any problems with the assumptions above?



Solution: From the scatterplot, we can see that when the return for X is above the mean μ_X, the return for Y tends to be above the mean μ_Y. Conversely, when the return for X is below μ_X, the return for Y tends to be below μ_Y. There seems to be a positive association between the two stocks. Thus, they are likely not independent, and the calculations in part (e), which assumed otherwise, are likely irrelevant.

Covariance

2. Parts (a)–(f) show plots of 50 random variable (X, Y) pairs sampled from four different 2-dimensional distributions. Dashed lines indicate the expectations of X and Y. In each part, decide if the covariance between X and Y seems to be positive, negative, or negligible.



Solution: The covariance looks positive in (a) and (e); it looks negative in (b), (c), and (f). The covariance looks negligible in (d).

3. Suppose X and Y are random variables with $\operatorname{var}(X) = 4$, $\operatorname{var}(Y) = 3$, and $\operatorname{cov}(X, Y) = -2$.

(a) Find $\operatorname{var}(X + Y)$.

Solution:

$$\begin{aligned}\operatorname{var}(X+Y) &= \operatorname{var}(X) + 2\operatorname{cov}(X,Y) + \operatorname{var}(Y) \\ &= 4 + 2 \cdot (-2) + 3 \\ &= 3.\end{aligned}$$

(b) Find $\operatorname{var}(2X + 5Y)$.

Solution:

$$\begin{aligned}\operatorname{var}(2X+5Y) &= 2^2 \operatorname{var}(X) + 2 \cdot 2 \cdot 5 \operatorname{cov}(X,Y) + 5^2 \operatorname{var}(Y) \\ &= 4 \cdot 4 + 20 \cdot (-2) + 25 \cdot 3 \\ &= 51.\end{aligned}$$

(c) Find $\operatorname{var}(3X - Y)$.

Solution:

$$\begin{aligned}\operatorname{var}(3X-Y) &= \operatorname{var}(3X + (-1)Y) \\ &= 3^2 \operatorname{var}(X) + 2 \cdot 3 \cdot (-1) \operatorname{cov}(X,Y) + (-1)^2 \operatorname{var}(Y) \\ &= 9 \cdot 4 - 6 \cdot (-2) + 1 \cdot 3 \\ &= 51.\end{aligned}$$

4. Suppose X and Y are random variables with means $\mu_X = 10$, $\mu_y = 5$, standard deviations $\sigma_X = 2$, $\sigma_Y = 4$, and correlation $\rho_{XY} = -.40$. Find $\text{var}(X+Y)$.

Solution: We first compute the covariance between X and Y:

$$\begin{aligned} \text{cov}(X,Y) &= \sigma_X \sigma_Y \rho_{XY} \\ &= (2)(4)(-.4) \\ &= -3.2. \end{aligned}$$

Now, we compute

$$\begin{aligned} \text{var}(X+Y) &= \text{var}(X) + 2\,\text{cov}(X,Y) + \text{var}(Y) \\ &= (2)^2 + 2 \cdot (-3.2) + (4)^2 \\ &= 13.6. \end{aligned}$$

5. Suppose X and Y are random variables with means $\mu_X = -10$, $\mu_y = 3$, standard deviations $\sigma_X = 4$, $\sigma_Y = 1$, and correlation $\rho_{XY} = .50$. Find $\text{var}(X - 2Y)$.

Solution: We first compute the covariance between X and Y:

$$\begin{aligned} \text{cov}(X,Y) &= \sigma_X \sigma_Y \rho_{XY} \\ &= (4)(1)(.50) \\ &= 2. \end{aligned}$$

Now, we compute

$$\begin{aligned} \text{var}(X-2Y) &= \text{var}(X+(-2)Y) \\ &= \text{var}(X) + 2(1)(-2)\,\text{cov}(X,Y) + (-2)^2\,\text{var}(Y) \\ &= (4)^2 - 4(2) + 4(1)^2 \\ &= 4. \end{aligned}$$