Covariance and independence?

I read from my textbook that $\text{cov}(X, Y) = 0$ does not guarantee X and Y are independent. But if they are independent, their covariance must be 0. I could not think of any proper example yet; could someone provide one?

independence covariance

edited Jul 11 '11 at 22:28
chl ♦ 35.4k 6 111 234

asked Jul 9 '11 at 19:47
Flying pig 1,249 3 16 24

8 You might also enjoy a quick review of Anscombe's Quartet, which illustrates some of the many different ways in which a particular nonzero covariance can be realized by a bivariate dataset. – whuber ♦ Jul 11 '11 at 18:01

5 The thing to note is that the measure of covariance is a measure of linearity.. Calculating the covariance is answering the question 'Do the data form a straight line pattern?' If the data do follow a linear pattern, they are therefore dependent. BUT, this is only one way in which the data can be dependent. Its like asking 'Am I driving recklessly?' One question might be 'Are you travelling 25 mph over the speed limit?' But that isn't the only way to drive recklessly. Another question could be 'Are you drunk?' etc.. There is more than one way to drive recklessly. – Adam Feb 16 '12 at 4:13

The so- called measure of linearity gives a structure to the relationship. What is important that the relationship can be non-linear which is not uncommon. Generally, covariance is not zero, It is hypothetical.The covariance indicates the magnitude and not a ratio, – subhash c. davar Sep 24 '13 at 15:55

4 Answers

Easy example: Let X be a random variable that is -1 or $+1$ with probability 0.5. Then let Y be a random variable such that $Y = 0$ if $X = -1$, and Y is randomly -1 or $+1$ with probability 0.5 if $X = 1$.

Clearly X and Y are highly dependent (since knowing Y allows me to perfectly know X), but their covariance is zero: They both have zero mean, and

$$\begin{aligned}\mathbb{E}[XY] &= (-1) \cdot 0 \quad \cdot P(X=-1) \\ &+ 1 \quad \cdot 1 \quad \cdot P(X=1, Y=1) \\ &+ 1 \quad \cdot (-1) \cdot P(X=1, Y=-1) \\ &= 0.\end{aligned}$$

Or more generally, take any distribution $P(X)$ and any $P(Y|X)$ such that $P(Y = a|X) = P(Y = -a|X)$ for all X (i.e., a joint distribution that is symmetric around the x axis), and you will always have zero covariance. But you will have non-independence whenever $P(Y|X) \neq P(Y)$; i.e., the conditionals are not all equal to the marginal. Or ditto for symmetry around the y axis.

edited Jul 11 '11 at 18:05
whuber ♦ 127k 13 236 461

answered Jul 9 '11 at 21:03
jpillow 847 8 8

Here is the example I always give to the students. Take a random variable X with $EX = 0$ and $EX^3 = 0$, e.g. normal random variable with zero mean. Take $Y = X^2$. It is clear that X and Y are related, but

$$cov(X, Y) = EXY - EX \cdot EY = EX^3 = 0.$$

edited Feb 16 '12 at 8:23

answered Jul 15 '11 at 8:18
mpiktas 22.7k 4 43 92

I like that example too. As a particular case, a N(0,1) rv and a chi2(1) rv are uncorrelated. – ocram Jul 15 '11 at 8:21

1 +1 but as a minor nitpick, you do need to assume that $E[X^3]=0$ separately (it does not follow from the assumption of symmetry of the distribution or from $E[X]=0$), so that we don't have issues such as $E[X^3]$ working out to be of the form $\infty - \infty$. And I am queasy about @ocram's assertion that "**a** N(0,1) rv and **a** chi2(1) rv are uncorrelated." (emphasis added) Yes, $X \sim N(0,1)$ and $X^2 \sim \chi^2(1)$ are uncorrelated, but not **any** $N(0,1)$ and $\chi^2(1)$ random variables. – Dilip Sarwate Feb 16 '12 at 3:09

@DilipSarwate, thanks, I've edited my answer accordingly. When I wrote it I though about normal variables, for them zero third moment follows from zero mean. – mpiktas Feb 16 '12 at 8:25

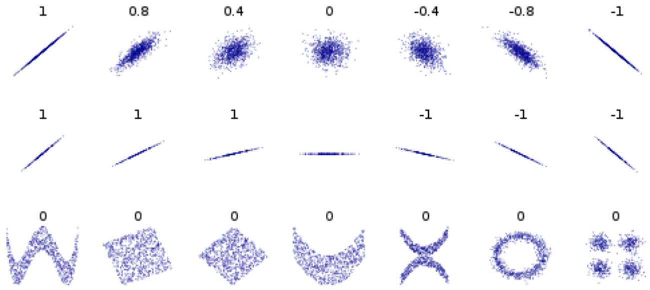
Some other examples, consider datapoints that form a circle or ellipse, the covariance is 0, but knowing x you narrow y to 2 values. Or data in a square or rectangle. Also data that forms an X or a V or a ^ or < or > will all give covariance 0, but are not independent. If y = sin(x) (or cos) and x covers an integer multiple of periods then cov will equal 0, but knowing x you know y or at least |y| in the ellipse, x, <, and > cases.

edited Jul 14 '11 at 18:13

answered Jul 11 '11 at 16:54
Greg Snow
31.1k 39 98

1 That if should be "if x covers an integer multiple of periods beginning at a peak or trough", or more generally: "If x covers an interval on which y is symmetric" – naught101 Feb 16 '12 at 0:47

The image below (source Wikipedia) has a number of examples on the third row, in particular the first and the fourth example have a strong dependent relationship, but 0 correlation (and 0 covariance).



edited Jan 21 '15 at 8:48
Tim
11.8k 2 28 61

answered Feb 16 '12 at 0:52
naught101
1,666 23 51