Does the variance of a sum equal the sum of the variances?

Asked 7 years, 3 months ago Active 3 years, 4 months ago Viewed 125k times

60

51

Is it (always) true that

$$\mathrm{Var}\left(\sum_{i=1}^{m} X_i\right) = \sum_{i=1}^{m} \mathrm{Var}(X_i)\ ?$$

variance

edited Jun 27 '12 at 2:30
Macro
34k 6 129 142

asked Jun 26 '12 at 22:44
Abe
1,613 5 22 39

3 The answers below provide the proof. The intuition can be seen in the simple case var(x+y): if x and y are positively correlated, both will tend to be large/small together, increasing total variation. If they are negatively correlated, they will tend to cancel each other, decreasing total variation. – Assad Ebrahim Jul 17 '15 at 10:51

4 Answers

89

The answer to your question is "Sometimes, but not in general".

To see this let $X_1, \dots, X_n$ be random variables (with finite variances). Then,

$$\mathrm{var}\left(\sum_{i=1}^{n} X_i\right) = E\left(\left[\sum_{i=1}^{n} X_i\right]^2\right) - \left[E\left(\sum_{i=1}^{n} X_i\right)\right]^2$$

Now note that $(\sum_{i=1}^n a_i)^2 = \sum_{i=1}^n \sum_{j=1}^n a_i a_j$, which is clear if you think about what you're doing when you calculate $(a_1 + \ldots + a_n) \cdot (a_1 + \ldots + a_n)$ by hand. Therefore, similarly,

$$\left[E\left(\sum_{i=1}^{n} X_i\right)\right]^2 = \left[\sum_{i=1}^{n} E(X_i)\right]^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} E(X_i)E(X_j)$$

so

$$\text{var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(E(X_iX_j) - E(X_i)E(X_j)\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\text{cov}(X_i, X_j)$$

by the definition of covariance.

Now regarding *Does the variance of a sum equal the sum of the variances?*:

- **If the variables are uncorrelated, yes**: that is, $\text{cov}(X_i, X_j) = 0$ for $i \neq j$, then

$$\text{var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\text{cov}(X_i, X_j) = \sum_{i=1}^{n}\text{cov}(X_i, X_i) = \sum_{i=1}^{n}\text{var}(X_i)$$

- **If the variables are correlated, no, not in general**: For example, suppose X_1, X_2 are two random variables each with variance σ^2 and $\text{cov}(X_1, X_2) = \rho$ where $0 < \rho < \sigma^2$. Then $\text{var}(X_1 + X_2) = 2(\sigma^2 + \rho) \neq 2\sigma^2$, so the identity fails.
- **but it is possible for certain examples**: Suppose X_1, X_2, X_3 have covariance matrix

$$\begin{pmatrix} 1 & 0.4 & -0.6 \\ 0.4 & 1 & 0.2 \\ -0.6 & 0.2 & 1 \end{pmatrix}$$

then $\text{var}(X_1 + X_2 + X_3) = 3 = \text{var}(X_1) + \text{var}(X_2) + \text{var}(X_3)$

Therefore **if the variables are uncorrelated** then the variance of the sum is the sum of the variances, but converse is **not** true in general.

edited Jun 28 '12 at 1:12

answered Jun 26 '12 at 22:51
Macro
34k 6 129 142

Regarding the example covariance matrix, is the following correct: the symmetry between the upper right and lower left triangles reflects the fact that $\text{cov}(X_i, X_j) = \text{cov}(X_j, X_i)$, but the symmetry between the upper left and the lower right (in this case that $\text{cov}(X_1, X_2) = \text{cov}(X_2, X_3) = 0.3$ is just part of the example, but could be replaced with two different numbers that sum to 0.6 e.g., $\text{cov}(X_1, X_2) = a$ and $\text{cov}(X_2, X, 3) = 0.6 - a$? Thanks again. – Abe Jun 27 '12 at 17:56

41

$$\mathrm{Var}\left(\sum_{i=1}^{m} X_i\right) = \sum_{i=1}^{m} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

So, if the covariances average to 0, which would be a consequence if the variables are pairwise uncorrelated or if they are independent, then the variance of the sum is the sum of the variances.

An example where this is not true: Let $\mathrm{Var}(X_1) = 1$. Let $X_2 = X_1$. Then $\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(2X_1) = 4$.

answered Jun 26 '12 at 22:59
Douglas Zare
9,363 1 31 43

It will rarely be true for sample variances. – DWin Jun 27 '12 at 2:35

1 @DWin, "rare" is an understatement - if the Xs have a continuous distribution, the probability that the sample variance of the sum is equal to the sum of the sample variances in exactly 0 :) – Macro Jun 27 '12 at 13:41

15

I just wanted to add a more succinct version of the proof given by Macro, so it's easier to see what's going on.

Notice that since $\mathrm{Var}(X) = \mathrm{Cov}(X, X)$

For any two random variables X, Y we have:

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathrm{Cov}(X+Y, X+Y) \\ &= E((X+Y)^2) - E(X+Y)E(X+Y) \\ &\text{by expanding,} \\ &= E(X^2) - (E(X))^2 + E(Y^2) - (E(Y))^2 + 2(E(XY) - E(X)E(Y)) \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2(E(XY)) - E(X)E(Y))\end{aligned}$$

Therefore in general, the variance of the sum of two random variables is not the sum of the variances. However, if X, Y are independent, then $E(XY) = E(X)E(Y)$, and we have $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Notice that we can produce the result for the sum of n random variables by a simple induction.

edited May 15 '16 at 15:45
Greenparker

answered May 15 '16 at 15:18
Omar Haque

11

Yes, if each pair of the X_i's are uncorrelated, this is true.

See the explanation on Wikipedia

edited Jun 27 '12 at 11:21

community wiki
3 revs, 3 users 67%
Abe

I agree. You also find a simple(r) explanation on Insight Things. – Jan Rothkegel Mar 11 '16 at 13:17