Variance

The **variance** is a single-valued metric that reflects the amount of spread that the values of a random variable will take on. More specifically, variance is the expected difference between the random variable's value and its mean (squared):

$$\text{Var}(X) := E\left[\ (X - E(X))^2\ \right]$$

(Definition 1). Variance can also be expressed as

$$\text{Var}(X) = E(X^2) - E(X)^2$$

as proven in Theorem 1.

Intuition

The variance of a random variable is single number that tells us about the amount of spread that we would expect to see if we were able to repeatedly sample from random variable's distribution. We note that the expectation of a random variable only tells us the average value of the random variable over a long number of observations, but it doesn't tell us anything about how spread out we expect these values to be. For example, let us define a random variable X where,

$$P(X = 0) = 0.5$$
$$P(X = 100) = 0.5$$

then,

$$E(X) = 50$$

If for another random variable,

$$P(Y = 50) = 1$$

then,

$$E(Y) = 50$$

Despite the fact that the two random variables behave very differently, they have the same expected value. The expected value didn't at all capture the fact that the values of X are much more spread out than Y's.

Properties

1. **Variance of a scaled random variable:**

   $$\text{Var}(cX) = c^2\text{Var}(X)$$

   where c is a constant (Theorem 2). Unlike expectation, variance is not a linear function.

2. **Variance of a shifted random variable:** Given a random variable X and constant c, the variance of $X + c$ is simply the variance of X:

$$\text{Var}(X + c) = \text{Var}(X)$$

(Theorem 3). This result makes intuitive sense; since the variance measures the amount of spread of the distribution, shifting the distribution left or right by a constant doesn't affect that spread and therefore shouldn't affect the variance.

3. **Variance of a point mass random variable:** Given a random variable for which $X = c$ with probability 1, the variance of X is zero. Furthermore, if X is not constant, then its variance is greater than zero (Theorem 4). This makes intuitive sense, if a random variable will always be the same value, then there is zero spread in the outcomes. On the other hand, if the random variable can take on more than one value (even with small probability), the average spread will be non-zero.

4. **Variance of convolution of random variables:** Given two random variables X and Y, variance of their sum is:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

(Theorem 5). If, however, X and Y are independent, then we have

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

since the covariance of independent random variables is always zero.

Definition 1 *Given random variable X with defined expected value, it's* ***variance*** *is given by*

$$Var(X) := E\left[\ (X - E(X))^2\ \right]$$

Theorem 1

$$Var(X) = E(X^2) - E(X)^2$$

Proof:

$$\begin{aligned}\text{Var}(X) &= E\left[\ (X - E(X))^2\ \right] \\ &= E\left[(X - E(X))(X - E(X))\right] \\ &= E\left[X^2 - 2XE(X) + E(X)^2\right] \\ &= E(X^2) - E[2XE(X)] + E[E(X)^2] \\ &= E(X^2) - 2E(X)E(X) + E(X)^2 \\ &= E(X^2) - 2E(X)^2 + E(X)^2 \\ &= E(X^2) - E(X)^2\end{aligned}$$

□

Theorem 2

$$Var(cX) = c^2 Var(X)$$

Proof:

$$\begin{aligned}\text{Var}(cX) &= E\left[(cX)^2\right] - E(cX)^2 \\ &= c^2 E\left(X^2\right) - c^2 E(X)^2 \\ &= c^2[E\left(X^2\right) - E(X)^2] \\ &= c^2 \text{Var}(X)\end{aligned}$$

□

Theorem 3

$$Var(X + c) = Var(X)$$

Proof:

$$\begin{aligned}\text{Var}(X + c) &= E\left[(X + c)^2\right] - E(X + c)^2 \\ &= E\left(X^2 + 2cX + c^2\right) - E(X + c)E(X + c)\end{aligned}$$

Expanding the first term,

$$E\left(X^2 + 2cX + c^2\right) = E\left(X^2\right) + 2cE(X) + c^2$$

Expanding the second term,

$$\begin{aligned} E(X+c)E(X+c) &= E(X)E(X+c) + E(c)E(X+c) \\ &= E[XE(X) + cE(X)] + cE(X) + cE(c) \\ &= E(X)^2 + cE(X) + cE(X) + c^2 \\ &= E(X)^2 + 2cE(X) + c^2 \end{aligned}$$

Putting it all together,

$$\begin{aligned} \text{Var}(X+c) &= E\left(X^2\right) + 2cE(X) + c^2 - E(X)^2 - 2cE(X) - c^2 \\ &= E\left(X^2\right) - E(X)^2 \\ &= \text{Var}(X) \end{aligned}$$

□

Theorem 4 *If a random variable X is equal to a constant c, then*

$$Var(X) = 0$$

Otherswise,

$$Var(X) \geq 0$$

Proof:

The proof of this property lies in the fact that variance is equal to $E\left[(X - E(X))^2\right]$. Not that if $X = c$, then the value inside the expectation is zero. Otherwise, the value inside the exception is positive due to the squared.

□

Theorem 5

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

and

$$Var(X-Y) = Var(X) + Var(Y) - 2Cov(X,Y)$$

Proof:

We only prove the result for $X+Y$. The result for $X-Y$ can be proven by identical calculation by substitution $-Y$ for Y.

$$\text{Var}(X+Y) = E\left[(X+Y)^2\right] - E(X-Y)^2$$

We see that the first term can be expressed as,

$$\begin{aligned} E\left[(X+Y)^2\right] &= E\left(X^2 + 2XY + Y^2\right) \\ &= E\left(X^2\right) + 2E(XY) + E\left(Y^2\right) \end{aligned}$$

The second term can be expressed as,

$$\begin{aligned} E(X+Y)^2 &= E(X+Y)E(X+Y) \\ &= E(X)E(X+Y) + E(Y)E(X+Y) \\ &= E[E(X)(X+Y)] + E[E(Y)(X+Y)] \\ &= E[XE(X) + XE(Y)] + E[XE(Y) + YE(Y)] \\ &= E[XE(X)] + E[XE(Y)] + E[XE(Y)] + E[YE(Y)] \\ &= E(X)^2 + E(X)E(Y) + E(X)E(Y) + E(Y)^2 \\ &= E(X)^2 + 2E(X)E(Y) + E(Y)^2 \end{aligned}$$

Now putting it all together,

$$\begin{aligned} \text{Var}(X+Y) &= E\left(X^2\right) + 2E(XY) + E\left(Y^2\right) - E(X)^2 - 2E(X)E(Y) - E(Y)^2 \\ &= E\left(X^2\right) - E(X)^2 + E\left(Y^2\right) - E(Y)^2 + 2[E(XY) - E(X)E(Y)] \\ &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y) \end{aligned}$$

□