planetmath.org

Math for the people, by the people.

variance

Canonical name	Variance
Date of creation	2013-03-22 11:53:46
Last modified on	2013-03-22 11:53:46
Owner	stevecheng (10074)
Last modified by	stevecheng (10074)
Numerical id	14
Author	stevecheng (10074)
Entry type	Definition
Classification	msc 62-00
Classification	msc 60-00
Classification	msc 81-00
Classification	msc 83-00
Classification	msc 82-00
Classification	msc 55-00
Related topic	GeometricDistribution2
Related topic	StandardDeviation
Related topic	Covariance
Related topic	MeanSquareDeviation

Definition

The *variance* of a real-valued random variable X is

$$\operatorname{Var} X = \mathbb{E}\left[(X-m)^2\right], \quad m = \mathbb{E}X,$$

provided that both expectations $\mathbb{E}X$ and $\mathbb{E}[(X-m)^2]$ exist.

The variance of X is often denoted by $\sigma^2(X)$, σ_X^2, or simply σ^2. The exponent on σ is put there so that the number $\sigma = \sqrt{\sigma^2}$ is measured in the same units as the random variable X itself.

The quantity $\sigma = \sqrt{\operatorname{Var} X}$ is called the *standard deviation* of X; because of the compatibility of the measuring units, standard deviation is usually the quantity that is quoted to describe an emprical probability distribution, rather than the variance.

Usage

The variance is a measure of the dispersion or variation of a random variable about its mean m.

It is not always the best measure of dispersion for all random variables, but compared to other measures, such as the absolute mean deviation, $\mathbb{E}[|X-m|]$, the variance is the most tractable analytically.

The variance is closely related to the $\mathbf{L}^2$ norm for random variables over a probability space.

Properties

1. The variance of X is the second moment of X minus the square of the first moment:
$$\operatorname{Var} X = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$
This formula is often used to calculate variance analytically.

2. Variance is not a linear function. It scales quadratically, and is not affected by shifts in the mean of the distribution:
$$\operatorname{Var}[aX+b] = a^2 \operatorname{Var} X, \quad \text{for any } a, b \in \mathbb{R}.$$

3. A random variable X is constant almost surely if and only if $\operatorname{Var} X = 0$.

4. The variance can also be characterized as the minimum of expected squared deviation of a random variable from any point:

$$\operatorname{Var} X = \inf_{a \in \mathbb{R}} \mathbb{E}[(X - a)^2]\,.$$

5. For any two random variables X and Y whose variances exist, the variance of the linear combination $aX + bY$ can be expressed in terms of their covariance:

$$\operatorname{Var}[aX + bY] = a^2 \operatorname{Var} X + b^2 \operatorname{Var} Y + 2ab \operatorname{Cov}[X, Y]\,,$$

where $\operatorname{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]$, and $a, b \in \mathbb{R}$.

6. For a random variable X, with actual observations $x_1, \ldots, x_n$, its variance is often estimated empirically with the *sample variance*:

$$\operatorname{Var} X \approx s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2\,, \quad \bar{x} = \frac{1}{n} \sum_{j=1}^{n} x_j\,.$$