Expected value, variance, standard deviation

Tom Coleman

Summary

This guide introduces expected values, variance, and standard deviation. These are key ideas in the concept of random variables, and they give information about the distribution of data conforming to a random variable.

Before reading this guide, it is highly recommended that you read Guide: PMFs, PDFs, and CDFs and [Guide: Introduction to data analysis].

Introduction

In Guide: PMFs, PDFs, and CDFs, you saw how these concepts are key tools in the study of probability, used to model and analyze the behaviour of random variables. As you have seen, these functions describe how probabilities are distributed across the possible outcomes of random events. For instance, the shape and position of the normal distribution (see Example 5 of Guide: PMFs, PDFs, and CDFs, or Factsheet: The normal distribution) can be modified by specifying the **mean** μ and the **standard deviation** σ. But what are these quantities? What do they mean, and how can they be computed from a given PMF or PDF?

This guide illustrates the related concepts of the **expected value**, **variance**, and **standard deviation** of a random variable X, and explains their usage and properties in probability theory. These concepts are not only found in statistics alone, but also in fields like classical and quantum mechanics, machine learning, and in decision theory.

Expected value

Often, the arithmetic mean of the outcomes is not the correct way to go when it comes to thinking about where the 'middle' of the set of data (conforming to a random variable). This is because some outcomes are more likely to occur than others. More formally, in a PMF or PDF describing a probability distribution, it is not guaranteed that every value is equally likely to occur.

For instance, suppose you have two six-sided dice; one is fair, and one is biased to give a lower value more often than not.

x	1	2	3	4	5	6
$\mathbb{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Table 1: PMF for rolling a fair six-sided die.

y	1	2	3	4	5	6
$\mathbb{P}(Y = y)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

Table 2: PMF for rolling a biased six-sided die.

You could ask; if you rolled one of these dice, what is the value that you would expect to see? With the fair, it would be between 3 and 4 as every outcome is equally likely to occur. With the biased die, you are more likely to get a smaller number as the probability of getting a number less than or equal to three is higher. This is because the PMFs for these dice are different.

So how can you measure the 'expected value' of a random variable given by some PMF or PDF? The answer is to 'weight' the outcomes by their probabilities; this is the definition of an **expected value**:

i Definition of the expected value

Suppose that you have a random variable X.

- If X is a discrete random variable with probability mass function $p(x)$, then the **expected value** $\mathbb{E}(X)$ of X is given by

$$\mathbb{E}(X) = \sum_x xp(x) = \sum_x x\mathbb{P}(X = x)$$

  that is, the sum of an outcome x multiplied by its probability $p(x) = \mathbb{P}(X = x)$ over all possible outcomes x.

- If X is a continuous random variable with probability density function $f(x)$, then the **expected value** $\mathbb{E}(X)$ of X is given by

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x$$

that is, the integral of x multiplied by its probability $f(x)$ over all possible values of x in the real numbers.

Tip

Geometrically, the expected value is the 'central point' of the distribution of the data. See below for more details on this.

Example 1

Let's look at the expected values of the two dice given above. The PMF for the fair die is given by

x	1	2	3	4	5	6
$\mathbb{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

so the expected value $\mathbb{E}(X)$ of this die is

$$\mathbb{E}(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6}$$

$$= \frac{1+2+3+4+5+6}{6} = 3.5$$

You can notice that this is the same as the mean of the outcomes; more on this later. What about for the second, biased die? This has PMF

y	1	2	3	4	5	6
$\mathbb{P}(Y = y)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

so the expected value $\mathbb{E}(Y)$ of this die is

$$\mathbb{E}(Y) = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{12} + 5 \cdot \frac{1}{12} + 6 \cdot \frac{1}{12}$$

$$= \frac{1+2+3}{4} + \frac{4+5+6}{12} = 2.75$$

which is **not** the mean of the outcomes.

Both of these 'expected values' are impossible to obtain in practice; try finding a die that shows 3.5 spots on the top! This is because the expected value actually represents the 'central point' of the distribution, rather than a specific outcome.

Tip

For a discrete random variable X with finitely many values, the expected value $\mathbb{E}(X)$ only equals the mean of the outcomes if and only if the probability of each outcome

$\mathbb{P}(X = x)$ is the same for every x. If this happens, then $\mathbb{P}(X = x) = 1/n$, and so the definition of expected value becomes

$$\mathbb{E}(X) = \sum_x x\mathbb{P}(X = x) = \sum_x x \cdot \frac{1}{n} = \frac{\sum_x x}{n}$$

which you can recognize as the mean.
But in all cases, the expected value of a random variable really is the 'population mean' of the probability distribution, which is not the same as the mean value of the outcomes. So **in theoretical probability distributions, the concept of the population mean of a random variable and the expected value of a random variable are the same.**

Here's an example of the expected value that introduces the idea of a **Bernoulli trial**.

Example 2

Imagine that you are in charge of an experiment that either succeeds with probability p or fails with probability $q = 1 - p$. You can write the outcome of success as 1 and the outcome of failure as 0. Therefore, the sample space is $\{0, 1\}$ with PMF $b(0) = q = 1 - p$ and $b(1) = p$. This is known as a **Bernoulli trial**.
Suppose that X is a random variable represented by a Bernoulli trial. What is the expected value $\mathbb{E}(X)$? You can use the definition of expected value to work this out:

$$\mathbb{E}(X) = 0 \cdot (1 - p) + 1 \cdot p = p$$

and so the expected value of a Bernoulli trial is the probability of success p.
Of course, much like the expected value of the dice above, this is impossible to obtain in practice!

Much like many other concepts in mathematics and statistics, you can use properties of expected values to build upon knowledge to find expected values of other random variables.

Properties of expected values

Let X be a random variable (either discrete or continuous). Then the following properties are hold for expected values:

1. If g is a function, then you can define a new random variable $Z = g(X)$ with the outcomes as $z = g(x)$.

- If X is discrete with PMF $f_X(x)$, then so is Z and

$$\mathbb{E}(Z) = \mathbb{E}(g(X)) = \sum_x g(x) f_X(x).$$

- If X is continuous with PDF $f_X(x)$, then so is Z and

$$\mathbb{E}(Z) = \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x.$$

From this and properties of summation and integration (see Guide: Sigma notation and [Guide: Properties of integration]) it follows that

2. If a, b are real number constants, then

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b.$$

3. If Y is another random variable of the same type as X (discrete or continuous), then

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

4. If Y is a random variable that is **independent** of X, then

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Tip

Properties 3 and 4 can be extended to any number of random variables (all pairwise independent in the case of 4.) For instance, if X_i are all random variables (of the same type) for $i = 1, \ldots, n$, then

$$\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i).$$

You can use the properties of the expected value to find the expected value of some more probability distributions.

Example 3

As seen in Guide: PMFs, PDFs, CDFs, a common example of a PMF is that of the **binomial distribution**. This is a type of PMF used to count the number of successes in a series of trials with only two possible outcomes: a success with probability p, or a failure with probability $q = 1 - p$. Here, the random variable X is 'number of successes'. Take x to be the number of successes in a number n of trials, p is the probability of success in a single trial, and $q = 1-p$ is the probability of failure. Then the PMF $f(x)$ for a binomial distribution is given by:

$$f(x) = \binom{n}{x} p^x q^{(n-x)}$$

How would you work out $\mathbb{E}(X)$ with this as the PMF? Instead, what you could do is use the fact that X is a random variable that encapsulates the repetition of n Bernoulli trials (see Example 2). So let Y be the random variable describing a Bernoulli trial; it follows that $X = nY$. You can then use Property 2 of the expected value to say that

$$\mathbb{E}(X) = \mathbb{E}(nY) = n\mathbb{E}(Y).$$

You know that $\mathbb{E}(Y) = p$ from Example 2, and so

$$\mathbb{E}(X) = n\mathbb{E}(Y) = np.$$

Variance and standard deviation

The expected value is the 'central point' of the probability distribution, which is only one of the key measures of the probability distribution. The other is how the probability is spread about this central point. This is known as the **variance** of the distribution. Related to the variance is the **standard deviation**.

Generally, the larger the variance, the more the data is spread out about the central point. Let's take a look at the definition of the variance:

Definition of variance and standard deviation

Suppose that you have a random variable X with expected value/mean $\mu = \mathbb{E}(X)$. The **variance** $\mathbb{V}(X)$ of X is given by

$$\mathbb{V}(X) = \mathbb{E}\left((X - \mu)^2\right)$$

that is, the expected value (average) of the squared deviation of the random variable X from the mean.
Using the fact that $\mu = \mathbb{E}(X)$ and properties of expected value, you can rephrase the variance by the formula

$$\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}\left(X\right)^2$$

The **standard deviation** σ is the square root of the variance; that is

$$\sigma = \sqrt{\mathbb{V}(X)}.$$

The standard deviation is so important that the variance is often written σ^2.

Tip

Much like sample means and population means, there is a concept of **sample variance** as well, to help line up distributions of samples compared to the overall distribution of the populations. See Guide: Introduction to hypothesis testing for more.

Example 4

What is the variance and standard deviation of the PMF for rolling a fair die? You could use the second formula for variance to work this out. You know from Example 1 that $\mathbb{E}(X) = 7/2 = 3.5$.

Now, you will need to work out $\mathbb{E}(X^2)$. You can use the PMF to write x^2 for every x, and then use this to work out $\mathbb{E}(X^2)$.

x	1	2	3	4	5	6
x^2	1	4	9	16	25	36

so $\mathbb{E}(X^2)$ is

$$\mathbb{E}(X^2) = 1 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} + 16 \cdot \frac{1}{6} + 25 \cdot \frac{1}{6} + 36 \cdot \frac{1}{6}$$

$$= \frac{1+4+9+16+25+36}{6} = \frac{91}{6}$$

You can now use this to work out the variance of the random variable:

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

$$= \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

Finally, the standard deviation is given by the square root of the variance, which means that $\sigma = \sqrt{35/12} \approx 1.708$.

Example 5

Consider a continuous random variable X uniformly distributed between a and b, which you have seen in Guide: PMFs, PDFs, CDFs. The PDF of X was given by:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

Let's find the expected value and variance of the uniform distribution.

Example 5 (continued)

Here, you can use the formula for the expected value (together with properties of integration, the definition of the probability density function, and the difference of two squares) to get

$$\begin{aligned}\mathbb{E}(X) &= \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{a} x \cdot 0\,\mathrm{d}x + \int_{a}^{b} \frac{x}{b-a}\,\mathrm{d}x + \int_{b}^{\infty} x \cdot 0\,\mathrm{d}x \\ &= 0 + \left[\frac{x^2}{2(b-a)}\right]_a^b + 0 = \frac{b^2-a^2}{2(b-a)} = \frac{1}{2}(a+b)\end{aligned}$$

This is what you would expect; it is the exact midpoint of the interval $[a, b]$!
To find the variance, you will need to find $\mathbb{E}(X^2)$. You can modify the above working to get

$$\begin{aligned}\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{a} x^2 \cdot 0\,\mathrm{d}x + \int_{a}^{b} \frac{x^2}{b-a}\,\mathrm{d}x + \int_{b}^{\infty} x^2 \cdot 0\,\mathrm{d}x \\ &= 0 + \left[\frac{x^3}{3(b-a)}\right]_a^b + 0 = \frac{b^3-a^3}{2(b-a)} = \frac{1}{3}(b^2+ab+a^2)\end{aligned}$$

Finally, you can say that

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \frac{1}{3}(b^2+ab+a^2) - \left(\frac{1}{2}(b+a)\right)^2 \\ &= \frac{1}{3}(b^2+ab+a^2) - \frac{1}{4}(b^2+2ab+a^2) \\ &= \frac{1}{12}(b^2-2ab+a^2) = \frac{1}{12}(b-a)^2\end{aligned}$$

and this is the variance.

Finally, the variance has properties like the expected value does:

Properties of variance

Let X be a random variable (either discrete or continuous). Then the following properties hold for the variance of X:

1. $\mathbb{V}(X) \geq 0$.
2. If a is a real number then
$$\mathbb{V}(aX) = a^2\mathbb{V}(X).$$
3. $\mathbb{V}(X + b) = \mathbb{V}(X)$ for all real numbers b.
4. If Y is another random variable of the same type as X (discrete or continuous), and Y is **independent** of X, then
$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y).$$

Tip

Property 4 can be extended to any number of pairwise independent random variables (all pairwise independent in the case of 4.) So if X_i are all pairwise independent random variables (of the same type) for $i = 1, \ldots, n$, then
$$\mathbb{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{V}(X_i).$$

Geometric interpretations of the expected value and variance

The figures below explore the normal distribution (for more, see Overview: Probability distributions) by changing the mean μ and the variance σ^2.

First of all, here is a figure of the normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$:

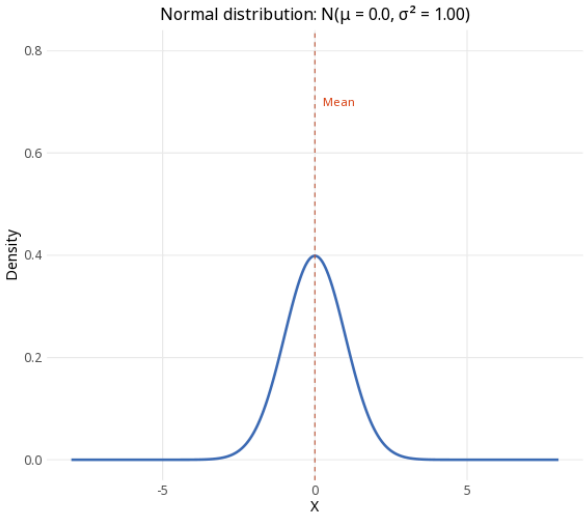

Figure 1: The normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$.

Next, here is a figure of the normal distribution with mean $\mu = -2$ and variance $\sigma^2 = 1$. You can see that changing the mean changes the centre of the curve.

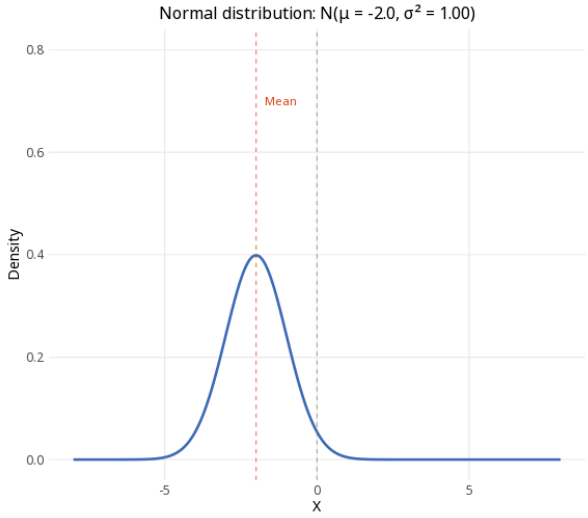

Figure 2: The normal distribution with mean $\mu = -2$ and variance $\sigma^2 = 1$.

Thirdly, here is a figure of the normal distribution with mean $\mu = -2$ and variance $\sigma^2 = 0.5$. You can see that decreasing the variance makes the curve far more narrow - as the data are more clustered toward the mean with a smaller variance.

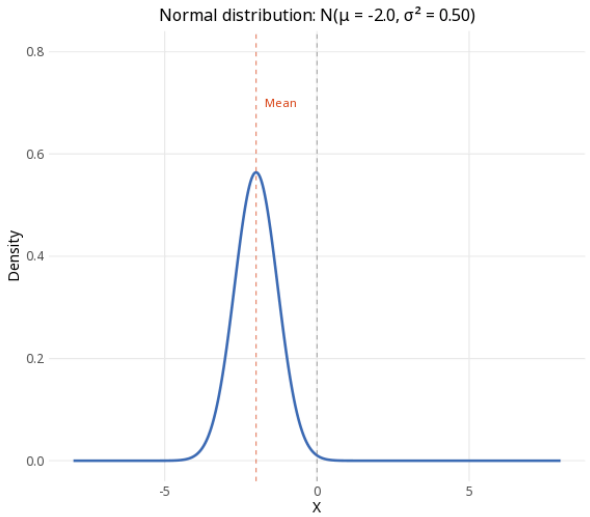

Figure 3: The normal distribution with mean $\mu = -2$ and variance $\sigma^2 = 0.5$.

Finally, here is a figure of the normal distribution with mean $\mu = -2$ and variance $\sigma^2 = 4$. You can see that increasing the variance leads to the curve being more spread out around the mean.

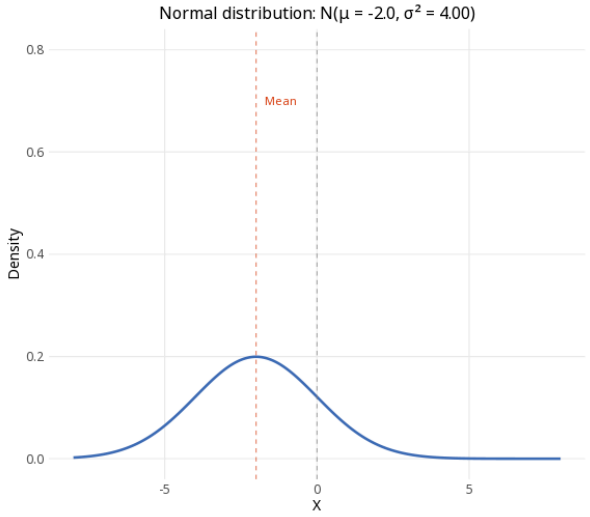

Figure 4: The normal distribution with mean $\mu = -2$ and variance $\sigma^2 = 4$.

You can see dynamic versions of these figures on the html version of this guide.

Quick check problems

1. Are the following statements true or false?

(a) The variance is a measure of how data is dispersed around the expected value:.

(b) The standard deviation is the square of the variance.

(c) The expected value is the 'central point' of the distribution.

2. Find the expected value and variance for rolling a fair 4-sided die.

3. Using the results from Examples 1 and 4 and the properties of expected values and variance, find the expected value and variance for rolling three fair 6-sided dice. You can assume that the events are independent of each other.

Further reading

For more questions on the subject, please go to Questions: Expected value, variance, standard deviation.

For more on why some PMFs and PDFs are valid, please go to Proof sheet: PMFs, PDFs, CDFs.

For more on probability distributions see Overview: Probability distributions.

Version history and licensing

v1.0: initial version created 08/25 by tdhc.

This work is licensed under CC BY-NC-SA 4.0.