Expectation value and Standard deviation

Suppose five students in class A took physics test and received following scores, which we call x_n^A:

$$x_1^A = 60, \qquad x_2^A = 70, \qquad x_3^A = 80, \qquad x_4^A = 95, \qquad x_5^A = 95 \tag{1}$$

What is $\langle x^A \rangle$, the average of their physics test scores? Since there are five students, it is given by following formula:

$$\langle x^A \rangle = \frac{60 + 70 + 80 + 95 \times 2}{5} = 80 \tag{2}$$

Now, suppose you choose one student at random. What is the probability that his or her score is 95? Since two students out of five received 95, we see that it is 2/5. How about 60? It is 1/5. For 70, it is 1/5, and for 80, it is also 1/5. Using these examples, we can express $\langle x^A \rangle$ slightly differently as follows:

$$\langle x^A \rangle = 60 \times \frac{1}{5} + 70 \times \frac{1}{5} + 80 \times \frac{1}{5} + 95 \times \frac{2}{5} \tag{3}$$

In other words, it is given by the sum of the scores multiplied by probability. This is called expectation value. It is the same concept as average, but the expression expectation value is more often used in the context of probability. Mathematically, we can write this as

$$\langle x \rangle = \sum_i x_i p_i \tag{4}$$

where there is a probability of p_i for the value x_i.

Now, suppose five students in class B took also physics test and received following scores, which we call x^B:

$$x_1^B = 75, \qquad x_2^B = 78, \qquad , x_3^B = 81, \qquad x_4^B = 81, \qquad x_5^B = 85 \tag{5}$$

As before, we can easily find that the average is 80. Does this mean that students at class B are as good as students at class A? Roughly speaking, yes, but precisely speaking, it is not that simple. Some students at class A are very good, while there are also some who are very bad. On the other hand, hardly any students at class B are good, but there aren't any that are so bad either. We can say that the scores for class B are more evenly distributed than the scores for class A. In other words, we can say that the scores of students at class A are more deviating from the average, then the ones at class B. Can we calculate "by how much?" Let's try. The deviations from average are given as follows:

$$x^A - \langle x^A \rangle = \qquad -20, \qquad -10, \qquad 0, \qquad 15, \qquad 15 \tag{6}$$

$$x^B - \langle x^B \rangle = \quad -5, \quad -2, \quad 1, \quad 1, \quad 5 \tag{7}$$

Let's sum them up to calculate the average of deviation. However, it is zero for all the both cases, as the negative deviations cancel the positive deviations, when they are summed. Actually, this is expected since:

$$\langle x^A - \langle x^A \rangle \rangle = \langle x^A \rangle - \langle x^A \rangle = 0 \tag{8}$$

And similarly for B. Here, we used the following facts:

$$\langle x - C \rangle = \langle x \rangle - C \tag{9}$$

where C is a number. In other words, the average of "something deducted by a certain number" is equal to "the average of the same thing" deducted by the same number. Actually, (8) can be equivalently derived more notationally complicatedly as follows. If we denote the number of students by N (i.e. in our case we have $N = 5$), this formula can be easily proven as follows:

$$\langle x^A - \langle x^A \rangle \rangle = \frac{\sum_{i=1}^{N} (x_i^A - \langle x^A \rangle)}{N} = \frac{(\sum_{i=1}^{5} x_i^A) - N \langle x^A \rangle}{N} \tag{10}$$

$$= \frac{\sum_{i=1}^{N} x_i^A}{N} - \frac{N \langle x^A \rangle}{N} = \langle x^A \rangle - \langle x^A \rangle = 0 \tag{11}$$

To make the negative deviations not cancel the positive deviations, the following trick turns out to be useful. We square the deviations and average them. This is called variance. (It actually turns out that this is much more useful than taking simply the absolute values. I didn't know this fact when I first studied statistics.) For example,

$$\mathrm{Var}(x^A) = \langle (x^A - \langle x^A \rangle)^2 \rangle = \frac{(-20)^2 + (-10)^2 + 0^2 + 15^2 + 15^2}{5} = 190 \tag{12}$$

We see that the squaring makes contributions for variance from negative deviations positive, as negative number squared is always positive. Now observe that this is an average of the value squared. So, it is meaningful to take a square root of it, to compare with the original value. This is called standard deviation, which we often denote by σ. In other words,

$$\sigma_A = \sqrt{\langle (x^A - \langle x^A \rangle)^2 \rangle} = \sqrt{190} \approx 13.8 \tag{13}$$

If we repeat the same calculation for class B, we obtain:

$$\sigma_B = \sqrt{\langle (x^B - \langle x^B \rangle)^2 \rangle} = \sqrt{11.2} \approx 3.3 \tag{14}$$

For class A, the students scores deviate from the average approximately by 13.8 roughly speaking. For class B, this is about 3.3. So, we conclude that the standard deviation of class B is much smaller than the standard deviation of class A. In other words, the scores for class B are more evenly distributed than the ones for class A.

Finally, we conclude this article with an important formula for variance. Readers will find it useful in "A short introduction to quantum mechanics XIII: Harmonic oscillator". Variance can be expressed as follows:

$$\begin{aligned} \mathrm{Var}(x) &= \langle (x - \langle x \rangle)^2 \rangle \\ &= \langle x^2 - 2\langle x \rangle x + \langle x \rangle \rangle^2 = \langle x^2 \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2 \\ &= \langle x^2 \rangle - \langle x \rangle^2 \end{aligned} \tag{15}$$

Here in the second line, we used the following fact:

$$\langle Dx \rangle = D\langle x \rangle \tag{16}$$

where D is a number. In other words, the average of "something multiplied by a certain number" is equal to "the average of the same thing" multiplied by the same number.

If you are not still sure how this is derived, here is another version of the proof:

$$\begin{aligned} \langle (x - \langle x \rangle)^2 \rangle &= \frac{\sum_i (x_i - \langle x \rangle)^2}{N} = \frac{\sum_i x_i^2 - 2\langle x \rangle x_i + \langle x \rangle^2}{N} \\ &= \frac{\sum_i x_i^2}{N} - 2\langle x \rangle \frac{\sum_i x_i}{N} + \frac{N\langle x \rangle^2}{N} \\ &= \langle x^2 \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2 = \langle x^2 \rangle - \langle x \rangle^2 \end{aligned} \tag{17}$$

Problem 1. Let's say that student A, B and C got 80, 90 and 100 on a math test respectively. What is the average of the score of these three students? What is the standard deviation?

Problem 2. Express $\langle (x - \langle x \rangle)^3 \rangle$ in terms of $\langle x \rangle$, $\langle x^2 \rangle$ and $\langle x^3 \rangle$.

Summary

- The expectation value of a score is given by the sum of the scores multiplied by probability. It is the same concept as average, but it is more often used in the context of probability. In other words,

  $$\langle x \rangle = \sum_i x_i p_i$$

  where there is a probability of p_i for x_i.

- The expectation value or the average of x is often denoted as $\langle x \rangle$.

- The variance of x is given by

  $$\mathrm{Var}(x) = \langle (x - \langle x \rangle)^2 \rangle$$

- The standard deviation is often denoted by σ and is given by the square root of the variance.

- $\mathrm{Var}(x) = \langle x^2 \rangle - \langle x \rangle^2$