Expectations

Santiago Yao

Sept 19, 2024

1 Definition

Roll a fair dice and you are the same likely to get 1, 2, 3, 4, 5, or 6. On "average", you expect to get $\frac{1+2+3+4+5+6}{6} = 3\frac{1}{2}$. The "average" outcome you get is called the **expectation**, written as $E[x]$ (mind the square brackets instead of a round pair), where x is the outcome of an **event**. Let's check out two more examples.

Example 1: Roll a fair dice with 1, 2, 3, 7, 8, 9 on each of its six faces. Record the number facing upwards and call it x. $E[x] = \frac{1+2+3+7+8+9}{6} = 5$.

Example 2: Toss a fair coin and it is the same likely to land head or tail. However, "head" and "tail" are not numbers and thus the expectation cannot be defined.

Now we have a more complicated example. A machine gives a random integer ranging from 1 to 4, and we call this number x. However, the machine is not fair. Indeed, $P(x = 1) = P(x = 2) = P(x = 3) = \frac{1}{6}$ and $P(x = 4) = \frac{1}{2}$. Now what is $E[x]$?

Needless to say, $E[x] = \frac{1+2+3+4}{4} = 2\frac{1}{2}$ would make a wrong response, since 4 carries more weight than other three numbers. Instead, 4 should be counted thrice as it carries thrice the weight than the other three. That is to say, we should image that we have numbers 1, 2, 3, 4, 4, 4 on a fair dice. Consequently, $E[x] = \frac{1+2+3+4+4+4}{6} = 3$.

The last example tells that, in essence, the expectation can be understood as the weighed average of all possible outcomes, where the "weight" is proportional to the probability of each outcome. Formally written,

$$E[x] = \sum_{x_i \in X} x_i * P(x_i),$$

where X denotes the set of all possible outcomes and it is guaranteed that $\sum_{x_i \in X} P(x_i) = 1$.[1]

Now, back to the last example. This time we apply the formal definition and it goes

$$E[x] = \sum_{x_i \in X} x_i P(x_i) = 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{2} = 3,$$

which is the same result as we attained just now.

Exercises

1. A fair dice with 4, 5, 6, 7, 8, 10 on each of its faces. Roll it once and the number facing upwards is recorded as x. Compute $E[x]$.
2. A fair dice with 4, 5, 6, 7, 8, a on each of its faces. Roll it once and the number facing upwards is recorded as x. Give $E[x] = 3$, what is a?
3. A machine gives a random number x from $\{3, 4, 5, 7\}$. The probabilities are as follows:$P(3) = \frac{1}{27}, P(4) = \frac{4}{27}, P(5) = \frac{20}{27}$ and $P(7)$ is untold. Compute $E[x]$.[2]
4. A teenage basketball club have 30 members of 14 years old, 90 members of 15, 75 members of 16, 30 members of 17, and 75 members of 18. What is the expectation of the age of a randomly picked member? (Here, "random" means that every member has the same chance of being picked.)

[1] If you need a brief introduction to the sigma notation, you may see mathcentre.ac.uk/resources/workbooks/mathcentre/sigma.pdf. By the way, $\in$ means "belongs to".

[2] Here $P(x = 3)$, etc. is written simply as $P(3)$, etc. This is the common practice whenever it does not hinder understanding.

2 Multiple events

Now consider this problem. Roll a fair dice of 1, 2, 3, 4, 5, and 6 on each of its faces twice. The outcome of the first time is x and that of the second time is y. Compute $E[x + y]$.

$x + y$ can take on any value ranging from 2 to 12. If we enumerate all possibilities, here is what we have got:

2 = 1 + 1, one case
3 = 1 + 2 = 2 + 1, two cases
4 = 1 + 3 = 2 + 2 = 3 + 1, three cases
5 = 1 + 4 = 2 + 3 = 3 + 2 = 4 + 1, four cases
6 = ..., five cases
7 = ..., six cases
8 = 2 + 6 = ... = 6 + 2, five cases (Pay attention here!)
9 = ..., four cases
10 = ..., three cases
11 = ..., two cases
12 = 6 + 6, one case

The total number of cases is 1+2+3+4+5+6+5+4+3+2+1 or 6×6 (why?), which is 36. Therefore,

$P(x + y = 2) = \frac{1}{36}$
$P(x + y = 3) = \frac{2}{36}$
$P(x + y = 4) = \frac{3}{36}$
$P(x + y = 5) = \frac{4}{36}$
$P(x + y = 6) = \frac{5}{36}$
$P(x + y = 7) = \frac{6}{36}$
$P(x + y = 8) = \frac{5}{36}$
$P(x + y = 9) = \frac{4}{36}$
$P(x + y = 10) = \frac{3}{36}$
$P(x + y = 11) = \frac{2}{36}$
$P(x + y = 12) = \frac{1}{36}$

Finally, $E[x+y] = \sum_{x+y=2}^{6}(x+y)P(x+y) = 2 * \frac{1}{36} + 3 * \frac{2}{36} + \cdots + 12 * \frac{1}{36} = 7$.

Generally speaking, when multiple events are concerned and you are asked to compute the expectation of certain expression that involves the outcome of every event, you should find the probability of each *final* outcome and then apply the definition ($E[x] = \sum_{x_i \in X} x_i * P(x_i)$). It requires you being rather careful and look after *all combinations of events*. There is usually no better way out.

In some special cases only, you can make your life less hard, which we will discuss in the next part.

Exercises
5. We have two sets: A = $\{2, 4, 6, 8\}$ and B = $\{1, 3, 5, 7\}$. Randomly pick x from A and y from B. Compute $E[x + y]$.
6. We have two sets: A = $\{5, 0, -5\}$ and B = $\{0, 3, 4\}$. Randomly pick x from A and y from B. Compute $E[x^2 + y^2]$.
7. We have three sets: A = $\{1,2\}$, P = $\{6, 7\}$, Q = $\{3, 4, 5\}$. First, one randomly pick a number from A and call it x. If x is 1, then one should turn to set P and randomly pick a number from it, and then call it y. If x is 2, then one should turn to set Q and randomly pick a number from it, and then call it y. Compute $E[xy]$.
8. A square has side length 1. Randomly pick any two distinct vertices of the square and record the distance between them as x. Another square has side length $\sqrt{2}$. Also randomly pick any two distinct vertices of the square and record the distance between them as y. Compute $E[xy]$.

3 Linearity

The linearity of expectations tells that

1. $E[x + y] = E[x] + E[y]$
2. $E[kx] = kE[x]$ for any k.

The proof of 2. is easy. Just notice that k is a constant that does not influence the outcome of x; it follows that $P(kx_i) = P(x_i)$. Thus, we have

$$E[kx] = \sum_{x_i \in x} kx_i P(kx_i) = \sum_{x_i \in x} kx_i P(x_i) = k \sum_{x_i \in x} x_i P(x_i) = kE[x]$$

.

Now turn to 1: $E[x+y] = E[x] + E[y]$. Let's say x has n outcomes $x_1, x_2, ..., x_n$ and y has m outcomes $y_1, y_2, ..., y_m$. Let $P_{i,j}$ be the possibility of combined outcome (x_i, y_j).

$$\begin{aligned}
E[x+y] &= \sum_{i,j} (x_i + y_j) P_{i,j} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} (x_i + y_j) P_{i,j} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i P_{i,j} + \sum_{i=1}^{n} \sum_{j=1}^{m} y_j P_{i,j} \\
&= \sum_{i=1}^{n} (x_i \sum_{j=1}^{m} P_{i,j}) + \sum_{j=1}^{m} (y_j \sum_{i=1}^{n} P_{i,j}) \\
&= \sum_{i=1}^{n} x_i P(x_i) + \sum_{j=1}^{m} y_j P(y_j) \\
&= E[x] + E[y]
\end{aligned}$$

It takes a while to make sense of the second last line above, only if you can see the simple fact that $P(x_i) = \sum_{i=1}^{n} P_{i,j}$ and $P(y_j) = \sum_{j=1}^{m} P_{i,j}$.

Now consider this problem once again: Roll a fair dice of 1, 2, 3, 4, 5, and 6 on each of its faces twice. The outcome of the first time is x and that of the second time is y. Compute $E[x+y]$.

$$\begin{aligned}
E[x+y] &= E[x] + E[y] \\
&= \frac{1+2+3+4+5+6}{6} + \frac{1+2+3+4+5+6}{6} \\
&= 7,
\end{aligned}$$

which is much easier than what we did in part 2.

When solving problems regarding linear combinations of several outcomes, you can use the linearity of expectation without proof.

Exercises

9. Solve exercise problem 5. once again, using the linearity of expectation.
10. Use exercise problem 7. to show that it is NOT true that $E[xy] = E[x]E[y]$.
11. Pick a random number from $\{1, 2, 3, 4, 5, 6\}$ twice. (Needless to say, the two numbers could be the same.) Call the first number x and the second y. Compute $E[3x + 5y]$.
12. (Leading to **The Bernoulli distribution**) Toss a fair coin 100 times. Each time, you receive 1 point if it lands head and 0 point if it lands tail. Your total score is x. What is $E[x]$?
13. A function $f(x)$ is defined on $\mathbf{Z}^+$ such that $f(x)$ is the greatest integer that is divisible by 3 and that does not exceed x. For example, $f(1) = 0, f(21) = 21, f(83) = 81$. Now, randomly choose x from $\{1, 2, 3, ..., 44\}$. Compute $E[f(x)] + E[f(x + 1)] + E[f(x + 2)]$.[3]

[3] $\mathbf{Z}^+$ means all positive integers.