

Expectation of Sum Proof

Now that we have learned about joint probabilities, we have all the tools we need to prove one of the most useful properties of Expectation: the fact that the expectation of a sum of random variables is equal to the sum of expectation (even if the variables are not independent). In other words:

For any two random variables X and Y ,

$$E[X + Y] = E[X] + E[Y]$$

The proof is going to use the Law of Unconscious statistician ([LOTUS](#)) where the function is addition!

Proof: Expectation of Sum

Let X and Y be any two random variables:

$$\begin{aligned} E[X + Y] &= \sum_x \sum_y (x + y) \cdot P(X = x, Y = y) && \text{LOTUS} \\ &= \sum_x \sum_y x \cdot P(X = x, Y = y) + y \cdot P(X = x, Y = y) && \text{Distribute} \\ &= \sum_x \sum_y x \cdot P(X = x, Y = y) + \sum_y \sum_x y \cdot P(X = x, Y = y) && \text{Rearrange Sums} \\ &= \sum_x x \sum_y P(X = x, Y = y) + \sum_y y \sum_x P(X = x, Y = y) && \text{Factor Out} \\ &= \sum_x x \cdot P(X = x) + \sum_y y \cdot P(Y = y) && \text{Def of Marginal} \\ &= E[X] + E[Y] && \text{Def of Expectation} \end{aligned}$$

At no point in the proof do we need to assume that X and Y are independent. In the second step the joint probability ends up in each sum, and in both cases, one of the sums ends up marginalizing over the joint probability!

Demonstration of the Proof

Here is an example to show the idea behind the proof. This table shows the joint probabilities $P(X = x, Y = y)$ for two random variables X and Y that are not independent. You will see how computing $E[X + Y]$ is the sum of terms that are used in $E[X]$ and $E[Y]$.

	$Y = 4$	$Y = 5$
$X = 1$	0.1	0.3
$X = 2$	0.2	0.4

Aside: These two random variables can each only take on two values. Having only four values in the joint table will make it easier to gain intuition.

Computing $E[X]$ using joint probabilities:

A key insight from the proof is that we can compute $E[X]$ using values from the joint. To do this we are going to use [marginalization](#):

$$P(X = x) = \sum_y P(X = x, Y = y)$$

We can expand $E[X]$ so that it is calculated only using values from the joint probability table:

$$\begin{aligned}
E[X] &= \sum_x x \cdot P(X = x) \\
&= \sum_x x \cdot \sum_y P(X = x, Y = y) && \text{Marginalization of } Y \\
&= \sum_x \sum_y x \cdot P(X = x, Y = y) && \text{Distribute } x
\end{aligned}$$

x	y	$P(X = x, Y = y)$	$x \cdot P(X = x, Y = y)$
1	4	0.1	$1 \times 0.1 = 0.1$
1	5	0.3	$1 \times 0.3 = 0.3$
2	4	0.2	$2 \times 0.2 = 0.4$
2	5	0.4	$2 \times 0.4 = 0.8$

$$E[X] = 0.1 + 0.3 + 0.4 + 0.8 = 1.6$$

Computing $E[Y]$ using joint probabilities:

Similarly, we can compute $E[Y]$ using only values from the joint:

$$\begin{aligned}
E[Y] &= \sum_y y \cdot P(Y = y) \\
&= \sum_y y \cdot \sum_x P(X = x, Y = y) && \text{Marginalization of } X \\
&= \sum_x \sum_y y \cdot P(X = x, Y = y) && \text{Distribute } y
\end{aligned}$$

x	y	$P(X = x, Y = y)$	$y \cdot P(X = x, Y = y)$
1	4	0.1	$4 \times 0.1 = 0.4$
1	5	0.3	$5 \times 0.3 = 1.5$
2	4	0.2	$4 \times 0.2 = 0.8$
2	5	0.4	$5 \times 0.4 = 2.0$

$$E[Y] = 0.4 + 1.5 + 0.8 + 2.0 = 4.7$$

Computing $E[X + Y]$ using joint probabilities:

We can rewrite $E[X + Y]$ to be the sum of terms used in the calculations of $E[X]$ and $E[Y]$ above:

$$\begin{aligned}
E[X + Y] &= \sum_{x,y} (x + y) \cdot P(X = x, Y = y) \\
&= \sum_{x,y} x \cdot P(X = x, Y = y) + y \cdot P(X = x, Y = y)
\end{aligned}$$

x	y	$P(x, y)$	$x \cdot P(x, y)$	$y \cdot P(x, y)$	$(x + y) \cdot P(x, y)$
1	4	0.1	0.1	0.4	$0.1 + 0.4 = 0.5$
1	5	0.3	0.3	1.5	$0.3 + 1.5 = 1.8$
2	4	0.2	0.4	0.8	$0.4 + 0.8 = 1.2$
2	5	0.4	0.8	2.0	$0.8 + 2.0 = 2.8$

Recall that $P(x, y)$ is shorthand for $P(X = x, Y = y)$.

Using the above derivation of the formula for $E[X + Y]$ in terms of values from the joint probability table:

$$E[X + Y] = \sum_{x,y} x \cdot P(X = x, Y = y) + y \cdot P(X = x, Y = y)$$

Plugging in values:

$$E[X + Y] = 0.1 + 0.4 + 0.3 + 1.5 + 0.4 + 0.8 + 0.8 + 2.0 = 6.3$$

We can observe that each of these values showed up exactly once when calculating $E[X]$ and $E[Y]$. This is why the proof works for any two random variables, even if they are not independent.

$$E[X] = 0.1 + 0.3 + 0.4 + 0.8 = 1.6$$

$$E[Y] = 0.4 + 1.5 + 0.8 + 2.0 = 4.7$$

Because they are summing the same values, it is no surprise that the sum of the expectations is equal to the expectation of the sum: $E[X + Y] = E[X] + E[Y] = 1.6 + 4.7 = 6.3$