Math 170E Week 7

Aidan Jan

February 24, 2023

Joint Distributions

Suppose X and Y are random variables. We want to know the distributions of X and Y. Notice this is not equal to knowing the two distributions separately, since X and Y may not be independent. (i.e., $\mathbb{P}(X = 2 \,\&\, Y = 3)$) In other words,

$$\forall k, l \, \mathbb{P}(X = k \& Y = l)$$

The Joint Probability Mass Function of X and Y would be a probability mass function distribution describing more than one variable occurring at a time.

$$\mathbb{P}_{X,Y}(k, l) := \mathbb{P}(X = k \,\&\, Y = l)$$

A joint pmf can be represented on a two-dimensional table, like the one shown below.

X, Y	5	6	8
2	0.1	0.2	0.1
7	0.15	0.05	0
8	0.3	0	0.1

Here are some example calculations using the table above.

$$\mathbb{P}(X = 2 \,\&\, Y = 8) = \mathbb{P}_{X,Y}(2, 8) = 0.1$$

$$\mathbb{P}(X < Y) = \mathbb{P}(X = 2 \,\&\, Y = 5) + \mathbb{P}(X = 2 \,\&\, Y = 6) + \mathbb{P}(X = 2 \,\&\, Y = 8) + \mathbb{P}(X = 7 \,\&\, Y = 8) = 0.4$$

$$\mathbb{P}(X = y) = 0.15 + 0.05 + 0 = 0.2$$

The final example calculation demonstrates the following Lemma:
Lemma:

$$\mathbb{P}_X(k) = \sum_l p_{X,Y}(k, l)$$

$$\mathbb{P}_Y(l) = \sum_k p_{X,Y}(k, l)$$

Also, like a normal, single variable pmf, the sum of all probabilities in a joint distribution must equal 1.

$$\sum_{k,l} p_{X,Y}(k, l) = 1$$

$$\sum_k p_X(k) = 1$$

An expected value can also be calculated.

$$\mathbb{E}\frac{X}{Y} = \sum_{k,l} \mathbb{P}(X = k \,\&\, Y = l) \cdot \frac{k}{l}$$

In the above example, this would equal:

$$0.1 \cdot \frac{2}{5} + 0.2 \cdot \frac{2}{6} + 0.1 \cdot \frac{2}{8} + \dots$$

The **general form** of expectation value of two variables is

$$\begin{aligned}\mathbb{E}f(X,Y) &= \sum_{k,l} \mathbb{P}(X = k \,\&\, Y = l) \cdot f(k,l) \\ &= \sum_{k,l} p_{X,Y}(k,l) \cdot f(k,l)\end{aligned}$$

This formula can also be extended to more than two variables:

$$\mathbb{E}f(X,Y,Z) = \sum_{k,l,m} p_{X,Y,Z}(k,l,m) \cdot f(k,l,m)$$

Returning to the Random Walk Problem

The original problem was: suppose you are on the middle of a number line from 1 to 10, inclusive. You are at 5 and every move, you randomly walk 1 number to the left or one number to the right. On average, how many moves would it take you to reach one of the ends?

Now, we will expand this problem. Consider the number line that is infinitely long, and you are at a random point. What is the probability that you will eventually reach 0 by randomly walking? This probability happens to be $\mathbb{P}(\text{eventually reach } 0) = 1$. This is because you have an infinite number of moves to traverse an infinite number of numbers. This converges to 1.

Now, what if instead of an infinite number line, we have an infinite plane? You start on a random (x, y) coordinate and want to go back to the point you started at. This probability also happens to be $\mathbb{P} = 1$, because you will eventually make it back in an infinite number of moves. However, the **expected value of the number of moves you take is infinite**. The probability you will make it back within an infinite number of moves is $\mathbb{P} = 1$, but the number of moves you are expected to take is $\mathbb{E} = \infty$.

Independence of Random Variables

Two random variables, X and Y are independent if

$$\forall k, l \; \mathbb{P}(Y = k | X = l) = \mathbb{P}(Y = k)$$

$$\forall k, l \; \mathbb{P}(X = l | Y = k) = \mathbb{P}(X = l)$$

If one is true, the other also will be true.
The statement

$$\forall k, l\, \mathbb{P}(X = k \,\&\, Y = l) = \mathbb{P}(X = k) \cdot \mathbb{P}(Y = l) = p_X(k) \cdot p_Y(l)$$

is also true. The formal definition goes as follows:

$$\boxed{X \perp Y \Leftrightarrow \forall k, l \in \mathbb{R}\{X = k\} \perp \{Y = l\}}$$

Note: The following is **NOT** true.

$$p_{X,Y}(k) = p_X(k) \cdot p_Y(k)$$

Identity formulas:

$$\mathbb{E}(f(x)|B) = \sum_k p_{X|B}(k) \cdot f(k)$$

$$\mathbb{E}f(x) = \sum_k p_X(k) \cdot f(k)$$

Lemmas:
Let $X \perp Y$. By the definition of perpendicular,

$$p_{X,Y} = p_X(k) \cdot p_Y(l) \forall k, l \in \mathbb{R}$$

1. $X \perp Y \implies \{X > 1\} \perp \{Y < 3\}$, or more specifically, $X \perp Y \implies \{X ****\} \perp \{Y ****\}$ Any event that is only dependent on one variable would also be independent of any event that is only dependent on the other variable.
Example:

$$\begin{aligned}\mathbb{P}(X > 1 \cap Y < 2) &= \mathbb{P}(X > 1) \cdot \mathbb{P}(Y < 2) \\ &= \left(\sum_{k>1} p_X(k)\right) \cdot \left(\sum_{l>2} p_Y(l)\right)\end{aligned}$$

2. $X \perp Y \implies f(x) \perp g(Y)$
Example:

$$\sin(X) \perp \cos(Y)$$

$$X^2 > 2 \perp Y^3 < 2$$

3. $\boxed{\mathbb{E}(X \cdot Y) = (\mathbb{E}X) \cdot (\mathbb{E}Y)}$ For this to apply, X and Y MUST be independent. However, $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ applies for all X and Y.
To prove that $\mathbb{E}(X \cdot Y) = (\mathbb{E}X) \cdot (\mathbb{E}Y)$, first write it in summation notation.

$$\mathbb{E}(X \cdot Y) = \sum_{k,l} p_{X,Y}(k, l) \cdot k \cdot l$$

However, $p_{X,Y} = p_X(k) \cdot p_Y(l)$. Therefore,

$$\begin{aligned}
&= \sum_{k,l} p_X(k) \cdot p_Y(l) \cdot k \cdot l \\
&= \left(\sum_k p_X(k) \cdot k\right) \cdot \left(\sum_l p_Y(l) \cdot l\right) \\
&= \mathbb{E}(X) \cdot \mathbb{E}(Y)
\end{aligned}$$

Going to the lemma $p_{X,Y}(k,l) = f(k) \cdot g(l) \forall k,l \implies X \perp Y$,
Example:

$$\begin{aligned}
p_{X,Y}(k,l) &= \begin{cases} \frac{1}{20} & \forall k \in [[1,4]], l \in [[1,5]] \\ 0 & \text{others} \end{cases} \\
&= f(k) \cdot g(l)
\end{aligned}$$

This is equal to

$$f(k) = \begin{cases} \frac{1}{20} & k \in [[1,4]] \\ 0 & \text{others} \end{cases}$$
$$g(l) = \begin{cases} 1 & l \in [[1,5]] \\ 0 & \text{others} \end{cases}$$

Using $\mathbb{E}(X \cdot Y) = \mathbb{E}X \cdot \mathbb{E}Y$, we can also prove that

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$$

Using the definition of variance,

$$\text{Var}(X \cdot Y) = \mathbb{E}((X+Y) - \mathbb{E}(X+Y))^2$$

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

$$\begin{aligned}
&= \mathbb{E}(X - \mathbb{E}(X) + Y - \mathbb{E}(Y))^2 \\
&= \mathbb{E}[(X - \mathbb{E}X)^2 + \mathbb{E}(2(X - \mathbb{E}X)(Y - \mathbb{E}Y)) + (Y - \mathbb{E}Y)^2 \\
&= [\text{Var}(X) + \mathbb{E}(2(X - \mathbb{E}X)(Y - \mathbb{E}Y)) + \text{Var}(Y)]
\end{aligned}$$

Therefore, as long as we prove $\mathbb{E}(2(X - \mathbb{E}X)(Y - \mathbb{E}Y)) = 0$, then we have proven the variance formula. First, we can pull the constant out.

$$\mathbb{E}(2(X - \mathbb{E}X)(Y - \mathbb{E}Y)) = 2 \cdot \mathbb{E}((X - \mathbb{E}X)(Y - \mathbb{E}Y))$$

Then, because $X \perp Y \implies f(X) \perp g(Y)$, the following is true:

$$(X - \mathbb{E}X) \perp (Y - \mathbb{E}Y)$$

Now, consider:

$$\mathbb{E}(X - \mathbb{E}X)$$

This can be rewritten to $\mathbb{E}X - \mathbb{E}(\mathbb{E}X)$ using a property of Expectation value. However, $\mathbb{E}(\mathbb{E}X) = \mathbb{E}X$, since the expected value of an expected value of a variable is just the expected value of a variable.
Thus,

$$\begin{aligned}\mathbb{E}(X - \mathbb{E}X) &= \mathbb{E}X - \mathbb{E}(\mathbb{E}X) \\ &= \mathbb{E}X - \mathbb{E}X \\ &= 0\end{aligned}$$

Q.E.D.

Linearity of Expectation Value

Going to the above example, the fact that Expectation Value is linear is very important. That is how the properties of Expectation Value, such as $\mathbb{E}(2X) = 2 \cdot \mathbb{E}(X)$ or $\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y$ are true. However, due to the linearity of Expectation Value, the following is also true:

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

Example:
Suppose you roll two die. Let X and Y be their values.
Let the event $B = \{X + Y = 7\}$.
What is $E(X|B)$ and $\mathbb{E}(Y|B)$?

It turns out that this is very easy to solve exploiting the linearity properties.
Since $E(X + Y|B) = \mathbb{E}(X|B) + \mathbb{E}(Y|B)$, we can add the two functions. Let $a = \mathbb{E}(X|X + Y = 7)$. Since X and Y both represent the independent outcomes of the dice, $a = \mathbb{E}(X|X + Y = 7)$. Thus,

$$2a = \mathbb{E}(X|B) + \mathbb{E}(Y|B)2a = \mathbb{E}(X + Y|B)$$

However, since $X+Y = 7$, as given by B, $2a = 7$. Thus, $a = \mathbb{E}(X|X+Y = 7) = \mathbb{E}(Y|X+Y = 7) = \frac{7}{2} = \boxed{3.5}$.