Lecture 15 : Pairs of Discrete Random Variables

Today we start Chapter 5. The transition we are making is like going from one variable calculus to vector calculus. We should really think of vectors (X, Y) of random variables.
So suppose X and Y are discrete random variables defined on the same sample space S.

Definition

The joint probability mass function, joint pmf, $P_{X,Y}(x, y)$, is defined by

$$P_{X,Y}(x, y) = P(X = x \overset{\substack{\text{and}\\ \downarrow}}{,} Y = y).$$

Example

A fair coin is tossed three times.

Let

$X = \sharp$ *of head on first toss*

$Y =$ *total* $\sharp$ *of heads*

As usual

$$S = \left\{ \begin{array}{llll} HHH, & HHT, & HTH, & HTT \\ THH, & THT, & TTH, & TTT \end{array} \right\}$$

We want to compute

$$\begin{aligned} P_{X,Y}(x,y) &= P(X = x, Y = y) \\ &= P((X = x) \cap (Y = y)) \\ &= P(X = x)P(Y = y | X = x) \end{aligned}$$

We will record the results in a matrix which we will now compute

$x \backslash y$	0	1	2	3
0				
1				

$P(X = 0, Y = 1)$

First column ($y = 0$)

Let's compute (upper left, $x = 0$)

$$P_{X,Y}(0,0) = P((X = 0) \cap (Y = 0))$$

$$= P(Y = 0) = P(TTT) = \frac{1}{8}$$

(because $Y = 0 \Rightarrow X = 0$)

Now lower left ($X = 1$)

$P_{X,Y}(1,0) = P(X = 1, Y = 0) = 0$

Move to the 2nd column ($y = 1$)

$P_{X,Y}(0,1) = P(X = 0, Y = 1)$ (top entry $X = 0$)

This is harder

$$P(X = 0, Y = 1) = P(T \text{ on first and exactly 1 head total})$$
$$= P(THT) + P(TTH) = \frac{2}{8}$$

The bottom entry of the second column is

$$P(X = 1, Y = 1) = P(HTT) = \frac{1}{8}$$

Third column $(y = 2)$

$$P(X = 0, Y = 2) = P(THH) = \frac{1}{8}$$
$$P(X = 1, Y = 2) = P(HTH) + P(HHT)$$
$$= \frac{2}{8}$$

Fourth column $(y = 3)$

$$P(X = 0, Y = 3) = 0$$
$$P(X = 1, Y = 3) = P(HHH) = \frac{1}{8}$$

The table for the joint *pmf* $P_{X,Y}(x, y)$

$X \backslash Y$	0	1	2	3
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

(*)

Check that the total probability is 1.
The joint *pmf* has a huge amount of information in it. In particular it contains the *pmf* $P_X(x)$ of X and $P_Y(y)$ of Y.
So how do we recover $P(Y = 1)$ from the table above. The event $(Y = 1)$ is the union of the two events $(X = 0, Y = 1)$ and $(X = 1, Y = 1)$. These two are mutually exclusive.

So

$$P(Y=1) = P(X=0, Y=1) + P(X=1, Y=1)$$
$$= \frac{2}{8} + \frac{1}{8} = \frac{3}{8}$$

= the sum of the entries in the second *column* (i.e. the column corresponding to $y = 1$)

How about $P(X = 1)$?

We have an equality of events

$$(X=1) = (X=1, Y=0) \cup (X=1, Y=1) \cup (X=1, Y=2) \cup (X=1, Y=3)$$
$$= 0 + \frac{1}{8} + \frac{2}{8} + \frac{1}{8} = \frac{1}{2}$$

$=$ the sum of the entries in the second *row* (corresponding to $X = 1$)

So we see we recover $P_Y(y)$ by taking *column sums* and $P_X(x)$ by taking *row sums*.

Marginal Distributions

We can express the above nicely by expanding the table (*) “adding margins”.

Table (*) with margins added

$x \backslash y$	0	1	2	3
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

(**)

The §64,000 question

How do you fill in the margins?
There is only one reasonable way to do this — put the row sums in the right margin and the column sums in the bottom margin.

Table (**) with the margins filled in

$x \backslash y$	0	1	2	3	
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0	$\frac{1}{2}$
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{1}{2}$
	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	

(***)

The right margin tells us the *pmf* of X and the bottom margin tells us the *pmf* of Y.

X	$\dots$	
0		$\frac{1}{2}$
1		$\frac{1}{2}$

y	0	1	2	3
$\vdots$				
	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

So we have

x	0	1
$P(X = x)$	$\frac{1}{2}$	$\frac{1}{2}$

and

y	0	1	2	3
$P(Y = y)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$X \sim \mathrm{Bin}\left(1, \frac{1}{2}\right) \qquad Y \sim \mathrm{Bin}\left(3, \frac{1}{2}\right)$$

For this reason, given the pair (X, Y) the *pmf*'s $P_X(x)$ and $P_Y(y)$ are called the marginal distributions.
To state all this correctly we have

Proposition

(i) $P_X(x) = \sum\limits_{all\ y} P_{X,Y}(X, y) \binom{row}{sum}$

(ii) $P_Y(y) = \sum\limits_{all\ x} P_{X,y}(x, y) \binom{column}{sum}$

So you "sum away" one variable leaving a function of the remaining variable.

Combining Discrete Random Variables

Suppose X and Y are discrete random variables defined on the same sample space. Let $h(x, y)$ be a real-valued function of two variables. We want to define a new *random variable* $W = h(X, Y)$.

Examples

We will start with the pair (X, Y) from our basic example.
The key point is that a function of a pair of random variables is again a random variable.

We will need only the joint *pmf*

$x \backslash y$	0	1	2	3
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

(*)

(i) $h(x, y) = x + y$ so $W = X + Y$

We see that the possible values of the sum are $0, 1, 2, 3, 4$ (since they are the sums of the possible values of X and Y).

We need to compute their probabilities. How do you compute

$$P(W = 0) = P(X + Y = 0)?$$

Answer

Find all the pairs x and y that add up to zero, take the probability of each such pair and add the resulting probabilities.

Answer (Cont.)

Bit $X + Y = 0 \Leftrightarrow X = 0$ *and* $Y = 0$ *so there is only one such pair* $(0, 0)$ *and (from the joint proof (*))*

$$P(X = 0, Y = 0) = \frac{1}{8}$$

Hence

$$P(W = 0) = P(X = 0, Y = 0) = \frac{1}{8}$$

$$\begin{aligned} P(W = 1) &= P(X + Y = 1) \\ &= P(X = 0, Y = 1) + P(X = 1, Y = 0) \\ &= \frac{2}{8} + 0 = \frac{2}{8} \end{aligned}$$

$$\begin{aligned} P(W = Z) &= P(X = 0, Y = Z) + P(X = 1, Y = 1) \\ &= \frac{2}{8} \end{aligned}$$

Answer (Cont.)

Similarly

$$P(W=3)=\frac{2}{8} \quad \text{and} \quad P(W=4)=\frac{1}{8}$$

So we get for $W = X + Y$

W	0	1	2	3	4
$P(W=w)$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

(b)

(check that the total probability is 1)

Remark

Technically the rule given in the "Answer" above is the definition of $W = X + Y$ *as a random variable but as usual the definition is forced on us.*

(ii) $h(X,y) = xy$ so $W = XY$

The possible values of W (the products of values of X with those of Y) are $0, 1, 2, 3$.

We now compute their probabilities.

P(W=0)

We can get 0 as a product xy if either $x = 0$ or $y = 0$ so we have

$$\begin{aligned}
P(W = 0) &= P(XY = 0) \\
&= P(X = 0, Y = 0) + P(X = 0, Y = 1) + P(X = 0, Y = 2) \\
&\quad + P(X = 0, Y = 3) + P(X = 1, Y = 0) \\
&= \frac{1}{8} + \frac{2}{8} + \frac{1}{8} + 0 + 0 = \frac{1}{2}
\end{aligned}$$

P(W=1)

$$P(W = 1) = P(X = 1, Y = 1) = \frac{1}{8}$$

$P(W = 2)$

$$P(W = 2) = P(X = 1, Y = 2) = \frac{2}{8}$$

$P(W = 3)$

$$P(W = 3) = P(X = 1, Y = 3) = \frac{1}{8}$$

W	0	1	2	3
$P(W = w)$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

(iii) $h(x, y) = \max(x, y) =$ the bigger of x and y
so $W = \max(X, Y)$

Remark

The max function doesn't turn up in vector calculus but it turns up a lot in statistics in advanced mathematics and real life.

The possible values of $\max(c, y)$ are $0, 1, 2, 3$.

$P(W = 0)$

$$\begin{aligned} P(W = 0) &= P(\mathrm{Max}(X, Y) = 0) \\ &= P(X = 0, Y = 0) = \frac{1}{8} \end{aligned}$$

$P(W = 1)$

$$\begin{aligned} P(W = 1) &= P(\mathrm{Max}(X, Y) = 1) \\ &= P(X = 0, Y = 1) + P(X = 1, Y = 0) \\ P(X = 1, Y = 1) &= \frac{3}{8} \end{aligned}$$

$P(W = 2)$

$$\begin{aligned} P(W = 1) &= P(X = 0, Y = 2) + P(X = 1, Y = 2) \\ &= \frac{3}{8} \end{aligned}$$

$$
\begin{aligned}
P(W = 3) &= P(X = 0, Y = 3) + P(X = 1, Y = 3) \\
&= \frac{1}{8}
\end{aligned}
$$

W	0	1	2	3
$P(W = w)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

(check that the total probability is 1)

The Expected Value of a Combination of Two Discrete Random Variables

If $W = h(X, Y)$ there are two ways to compute $E(W)$.

Proposition

$$E(W) = \sum_{\substack{\text{all } (x,y) \\ \text{possible} \\ \text{values of} \\ (X,Y)}} h(x, y) P_{X,Y}(x, y) \qquad (\sharp)$$

We will illustrate the proposition by computing $E(W)$ for the $W = X + Y$ of pages 12, 13, 14.

In *two ways?*

First way (without using the proposition)

W is a random variable with proof given by (b) on page 14.
(so we use (b))

$$
\begin{aligned}
E(W) &= (0)\left(\frac{1}{8}\right) + (1)\left(\frac{2}{8}\right) + (2)\left(\frac{2}{8}\right) \\
&\quad + (3)\left(\frac{2}{8}\right) + (4)\left(\frac{1}{8}\right) \\
&= \frac{2+4+6+4}{8} = \frac{16}{8} = 2
\end{aligned}
$$

Second way (using the proposition)

Now we use (*) from page 12

$$
E(W) = E(X+Y) = \underbrace{\sum_{\text{all } x,y} (x+y) P_{X,Y}(x,y)}_{\substack{\text{sum over the 8 entries} \\ \text{of (*)}}}
$$

$$
\begin{aligned}
&= (0+0)\left(\frac{1}{8}\right) + (0+1)\left(\frac{2}{8}\right) + (0+2)\left(\frac{1}{8}\right) + (0+3)(???) \\
&\quad (1+0)(0) + (1+1)\left(\frac{1}{8}\right) + (1+2)\left(\frac{2}{8}\right) + (1+3)\left(\frac{1}{8}\right) \\
&= \frac{2+2}{8} + \frac{2+6+4}{8} \\
&= \frac{4+12}{8} = 2
\end{aligned}
$$

The first way is easier but we need to compute the proof of $W = X + Y$ first. That was hard work, pages 12-14.