2.6.1. Joint Probability Mass Functions

Just as the pmf of a discrete random variable X must describe the probability of each value of X, the **joint pmf** of X and Y must describe the probability of any *combination* of values that X and Y could have. We could describe the joint pmf of random variables X and Y using a table like the one in the following example.

Example 2.6.1. The joint distribution of X and Y is described by the following table:

		value of X		
		1	2	3
	1	0.17	0.15	0.08
value of Y	2	0.00	0.10	0.10
	3	0.08	0.20	0.12

We can now calculate the probabilities of a number of events:

- $\mathrm{P}(X = 2) = 0.15 + 0.10 + 0.20 = 0.45$,
- $\mathrm{P}(Y = 2) = 0.00 + 0.10 + 0.10 = 0.20$,
- $\mathrm{P}(X = Y) = 0.17 + 0.10 + 0.12 = 0.39$,
- $\mathrm{P}(X > Y) = 0.15 + 0.08 + 0.10 = 0.33$.

The first two probabilities show that we can recover the pmfs for X and for Y from the joint distribution. These are known as the **marginal distributions** of X and Y because they can be obtained by summing across rows or down columns, and the natural place to record those values is in the margins of the table:

		value of X			
		1	2	3	total
	1	0.17	0.15	0.08	0.40
value of Y	2	0.00	0.10	0.10	0.20
	3	0.08	0.20	0.12	0.40
	total	0.25	0.45	0.30	1.00

We can also compute **conditional distributions**. Conditional distributions give probabilities for one random variable for a given value of the other. For example,

$$\mathrm{P}(Y = 2 \mid X = 2) = \frac{\mathrm{P}(Y = 2 \text{ and } X = 2)}{\mathrm{P}(X = 2)} = \frac{0.10}{0.45} = \frac{2}{9} = 0.2222 \,.$$

Notice the general form that applies here:

$$\text{conditional} = \frac{\text{joint}}{\text{marginal}} \,. \qquad \triangleleft$$

Tables like the ones in Example 2.6.1 are really just descriptions of functions, so our formal definitions are the following.

Definition 2.6.1. The *joint pmf* of a pair of discrete random variables $\langle X, Y \rangle$ is a function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that for all x and y,

$$\mathrm{P}(X = x \text{ and } Y = y) = f(x, y) \,. \qquad \square$$

Definition 2.6.2. If f is the joint pmf for X and Y, then the *marginal distributions* of X and Y are defined by

- $f_X(x) = \sum_y f(x,y)$ and
- $f_Y(y) = \sum_x f(x,y)$. □

Definition 2.6.3. If f is the joint pmf for X and Y, then the *conditional distributions* $f_{X|Y=y}$ and $f_{Y|X=x}$ are defined by

- $f_{X|Y=y}(x) = \dfrac{f(x,y)}{f_Y(y)}$ and
- $f_{Y|X=x}(y) = \dfrac{f(x,y)}{f_X(x)}$. □

An important (and generally easier) special case of joint distributions is the distribution of independent random variables.

Definition 2.6.4. Random variables X and Y are *independent* if for every x and y

$$f(x,y) = f_X(x) \cdot f_Y(y) \ ,$$

that is, if

$$\mathrm{P}(X = x \text{ and } Y = y) = \mathrm{P}(X = x) \cdot \mathrm{P}(Y = y) \ .$$ □

Example 2.6.2. The variables X and Y from Example 2.6.1 are not independent. This can be seen by observing that

- $\mathrm{P}(X = 2) = 0.45$ and
- $\mathrm{P}(Y = 2) = 0.20$, but
- $\mathrm{P}(X = 2 \text{ and } Y = 2) = 0.10 \neq 0.45 \cdot 0.20 = 0.09$

or that

- $\mathrm{P}(X = 1) = 0.25$ and
- $\mathrm{P}(Y = 2) = 0.20$, but
- $\mathrm{P}(X = 1 \text{ and } Y = 2) = 0.00 \neq 0.25 \cdot 0.20$.

The fact that

- $\mathrm{P}(X = 3) = 0.30$,
- $\mathrm{P}(Y = 3) = 0.40$, and
- $\mathrm{P}(X = 3 \text{ and } Y = 3) = 0.12 = 0.30 \cdot 0.40$

is not enough to make the variables independent. ◁

Just as with events, two random variables are independent if and only if the conditional distributions are equal to the marginal distributions.

Lemma 2.6.5. *X and Y are independent discrete random variables if and only if for all x and y*

(1) $f_{X|Y=y}(x) = f_X(x)$ *and*
(2) $f_{Y|X=x}(y) = f_Y(y)$.

Proof. Exercise 2.76. □

Example 2.6.3.

Q. Let $f(x,y) = \frac{x^2y}{84}$ on $\{1,2,3\}^2$. Show that f is a joint pmf and determine the marginal pmfs for X and Y. Are X and Y independent?

A.

$$\sum_{x,y} f(x,y)$$
$$= \frac{1}{84}\left(1^2\cdot 1 + 1^2\cdot 2 + 1^2\cdot 3 + 2^2\cdot 1 + 2^2\cdot 2 + 2^2\cdot 3 + 3^2\cdot 1 + 3^2\cdot 2 + 3^2\cdot 3\right)$$
$$= \frac{1}{84}(1+2+3+4+8+12+9+18+27) = 1\ ,$$

so f is a pmf.

The marginal distributions are

$$f_X(x) = f(x,1) + f(x,2) + f(x,3) = \frac{x^2(1+2+3)}{84} = \frac{6x^2}{84} = \frac{x^2}{14}$$

and

$$f_Y(y) = f(1,y) + f(2,y) + f(3,y) = \frac{(1+4+9)y}{84} = \frac{14x^2}{84} = \frac{6y}{6}$$

so

$$f_X(x)f_Y(y) = \frac{x^2}{14}\cdot\frac{y}{6} = \frac{x^2y}{84} = f(x,y)\ ,$$

which shows that X and Y are independent. ◁

2.6.2. Transformations

Just as we defined the transformation of a single random variable, we can now define functions of jointly distributed random variables.

Definition 2.6.6. Let f be the joint pmf of discrete random variables X and Y and let $t : \mathbb{R}\times\mathbb{R}\to\mathbb{R}$. Then $t(X,Y)$ is a discrete random variable with pmf given by

$$\mathrm{P}\left(t(X,Y)=k\right) = \sum_{t(x,y)=k} \mathrm{P}(X=x \text{ and } Y=y) = \sum_{t(x,y)=k} f(x,y)\ .$$ □

Example 2.6.4. Continuing with Example 2.6.1, we can determine the pmf for $X+Y$. The possible values of $X+Y$ are from 2 to 6. The table below is formed