stat 260 : only joint distributions

Set 18: Joint Probability Distributions

Stat 260: June 26, 2024

So far, we have looked at ***univariate*** *random variables. There are, however, many scenarios were several random variables interact with one another. For example, we might want to measure a subject's height X and weight Y; for this* ***multivariate*** *example, we would need* ***jointly distributed*** *random variables, and their* ***multivariate distributions****.*

Let X and Y be discrete random variables. The **joint probability mass function** (**joint pmf**) $f(x, y)$ is

$$f(x, y) = P(X = x \cap Y = y) = P(X = x, Y = y)$$

This joint pmf satisifies:

(i) $0 < f(x,y) \leq 1$

(ii) $\sum_{\text{all } y} \left(\sum_{\text{all } x} f(x,y) \right) = 1$

Example 1: Suppose that the discrete random variables X and Y have the following joint pmf:

	$f(x, y)$	Y = 3	6	9	
X	-2	0.12	0.09	0.10	0.31
	5	0.14	0.21	0.04	0.39
	7	0.06	0.11	0.13	0.30
	marginal for y →	0.32	0.41	0.27	↑ marginal for x

check → sum all probabilities inside =1

(a) $f(5, 3)$ $= P(X=5, Y=3) = 0.14$ from table

(b) $P(X = 5)$ $= f(5,3) + f(5,6) + f(5,9)$ → sum of all 5's or $\sum_{\text{all } y} f(5,y)$

$= 0.14 + 0.21 + 0.04 = 0.39$

(c) $P(Y \geq 6)$ $= P(Y=6) + P(Y=9)$

$P(Y=6) = 0.09 + 0.21 + 0.11 = 0.41$

$P(Y=9) = 0.10 + 0.04 + 0.13 = 0.27$

$= 0.68$ or you could've done $1 - P(Y=3)$

The ***marginal probability mass function distributions*** of **discrete** random variables X and Y are

$$\text{marginal for } x \rightarrow \quad f_X(x) = \sum_{\text{all } y} f(x,y) \quad \text{and} \quad f_Y(y) = \sum_{\text{all } x} f(x,y) \leftarrow \text{marginal for } y$$

Example 1 Continued...

(d) Determine the marginal pmf of X

x	-2	5	7
$f_x(x)$	0.31	0.39	0.3

← should sum to 1

(e) Find $E[X]$

$E(x) = (-2)(0.31) + 5(0.39) + 7(0.3) = 3.43$

Let X and Y be two random variables. The ***conditional probability distribution*** of the random variable Y **given** $X = x$ is

$$f_{Y|X}(y|x) = f_{Y|X=x}(y) = \boxed{\frac{f(x,y)}{f_X(x)}} = f(y=y \mid x=x)$$

Likewise, the conditional distribution of X given that $Y = y$ is

$$f_{X|Y}(x|y) = f_{X|Y=y}(x) = \frac{f(x,y)}{f_Y(y)}$$

Example 1 Continued...

(g) Determine $f_{Y|X}(9|7)$ → probability $y=9$ given $x=7$

$= P(y=9 \mid x=7)$

$\rightarrow \boxed{\frac{P(y=9 \cap x=7)}{P(x=7)}}$ they mean the same $\frac{f(7,9)}{f_x(7)}$

$\rightarrow \frac{0.13}{0.30}$ → y is 9 x is 7 from table on page 1

$\rightarrow 0.4333$

The random variables X and Y are **independent** if for every (x, y) pair,

$$\text{joint pmf} \rightarrow f(x,y) = f_X(x)f_Y(y). \leftarrow \text{marginals}$$

Example 1 Continued...

(f) Determine whether X and Y are independent.

Check : $f(5,3) = f_X(5) \cdot f_Y(3)$

$f(5,3) = 0.14$ ← not equal →

$f_X(5) \cdot f_Y(3) = (0.39) \cdot (0.32) = 0.1248$

$\therefore$ X and Y are independent ← since we found a value that is not equal

---

Note: *To determine whether the (discrete or continuous) random variables $X_1, X_2, \ldots, X_n$ are independent, we would have to test whether* every *subset $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ of variables (i.e. every pair, every triple, etc.) has the joint pmf of the subset equal to the product of the marginal pmf.*

> If the random variables $X_1, X_2, \ldots, X_n$ are independent, then
> $$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_n}(x_n)$$
> ↙ we can build the pmf from the marginals

Example 2: Suppose that the discrete random variables X and Y are independent with pmfs:

x	0	1	2	3
$f(x)$	0.3	0.1	0.4	0.2

y	2	4
$f(y)$	0.65	0.35

Find the joint pmf, $f(x, y)$.

y

$f(x,y)$ (x ↓)	2	4	
0	0.195	0.105	0.3
1	0.065	0.035	0.1
2	0.26	0.14	0.4
3	0.13	0.07	0.2
	0.65	0.35	

check all entries sum to 1

✳ Since X and Y are independent the marginals for $f(x,y)$ are just the same as the original distributions for X and Y.

$$\underbrace{f(0,2) = P(x=0) \cdot P(y=2)}_{\text{only true because we know X and Y are independent}}$$

$= (0.3) \cdot (0.65)$

$= 0.195$

Readings and Practice problems: See Set 19