Joint Distributions

Anastasiia Kim

April 6, 2020

Discrete Joint Distributions

When we build the joint distribution of two discrete random variables we can make a two-way table like this:

Condition/Gender	Male (M)	Female (F)	Total
Colorblind (C)	15/510	2/510	17/510
Not Colorblind (NC)	240/510	253/510	493/510
Total	1/2	1/2	1

$$P(C) = \frac{17}{510}$$

$$P(NC \text{ and } M) = \frac{240}{510}$$

$$P(C \text{ and } F) = \frac{2}{510}$$

Discrete Joint Distributions. Two random variables

For any two discrete random variables X and Y, the *joint probability mass function* of X and Y is

$$p(x, y) = P(X = x, Y = y)$$

The *joint pmf* satisfies

- $$p(x, y) \geq 0$$
- $$\sum_{\text{all } x} \sum_{\text{all } y} p(x, y) = 1$$

As usual, comma means 'and'

$$P(X = NC, Y = M) = \frac{240}{510}$$

$$P(X = C, Y = F) = \frac{2}{510}$$

Marginal PMFs

- The joint pmf contains all the information regarding the distributions of X and Y.
- We can obtain pmf of X from its joint pmf with Y

$$p(x) = P(X = x) = \sum_{\text{all } y} P(X = x, Y = y) = \sum_{\text{all } y} p(x, y)$$

  $p(x)$ is called the *marginal pmf* of X.
- the marginal pmf of Y is

$$p(y) = P(Y = y) = \sum_{\text{all } x} P(X = x, Y = y) = \sum_{\text{all } x} p(x, y)$$

The marginal probabilities in the example are

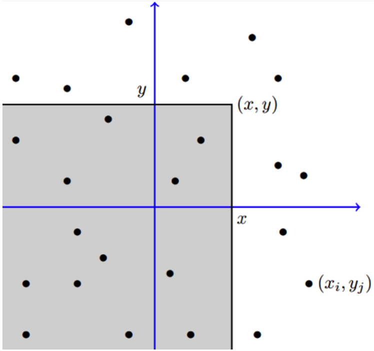
$$P(X = C) = P(X = C, Y = M) + P(X = C, Y = F) = \frac{15}{510} + \frac{2}{510} = \frac{17}{510}$$

$$P(Y = M) = \frac{15}{510} + \frac{240}{510} = \frac{255}{510} = \frac{1}{2}$$

Joint Cumulative Distributive Function

- Recall that, for a r.v. X, the cdf is $F(x) = P(X \leq x)$.
- For two r.v.s X and Y, the joint cdf is

$$F(x, y) = P(X \leq x, Y \leq y)$$



- The plot shows the shaded region associated with $F(x, y)$.
- Note that the above definition of joint cdf is a general definition and is applicable to discrete and continuous random variables.

Conditioning and Independence

- Recall that, the main formula for the conditional probability when $P(B) > 0$ is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- When two random variables are defined in a random experiment, knowledge of one can change the probabilities that we associate with the values of the other.
- In some cases, we have observed the value of a r.v. Y and need to update the pmf of another r.v. X whose value has not yet been observed. Use the *conditional pmf* of X given Y:

$$p_{X|Y}(x) = P(X = x|Y = y) = \frac{P(X = x \cap Y = y)}{P(Y = y)} = \frac{P(X = x, Y = y)}{P(Y = y)}$$

  Apply the formula for $P(A|B)$ with the event A defined to be $X = x$ and event B defined to be $Y = y$.

Conditional probability

Condition/Gender	Male (M)	Female (F)	Total
Colorblind (C)	15/510	2/510	17/510
Not Colorblind (NC)	240/510	253/510	493/510
Total	1/2	1/2	1

$$P(X = C|Y = M) = \frac{P(X = C, Y = M)}{P(Y = M)} = \frac{15/510}{1/2} = \frac{30}{510}$$

$$P(X = C|Y = F) = \frac{P(X = C, Y = F)}{P(Y = F)} = \frac{2/510}{1/2} = \frac{4}{510}$$

Note that

$$P(X = C|Y = M) + P(X = NC|Y = M) = 1$$

$$P(Y = M|X = C) + P(Y = F|X = C) = 1$$

Independent Random Variables

Two discrete random variables X and Y are independent if for all x,y

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

Equivalently, X and Y are independent if for all x,y

$$F(x, y) = F(x)F(y)$$

If X and Y are independent, we have

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{P(X = x)P(Y = y)}{P(Y = y)} = P(X = x)$$

which means that the conditional pmf is equal to the marginal pmf. In other words, knowing the value of Y does not provide any information about X.

Independence

	Y = 2	Y = 4	Y = 5
X = 1	1/12	1/24	1/24
X = 2	1/6	1/12	1/8
X = 3	1/4	1/8	1/12

- find $P(X \leq 1, Y \leq 4) = 1/12 + 1/24 = 1/8$
- the marginal pmf of X is

X = x	X = 1	X = 2	X = 3	otherwise
p(x)	1/6	3/8	11/24	0

- the marginal pmf of Y is

Y = y	Y = 2	Y = 4	Y = 5	otherwise
p(y)	1/2	1/4	1/4	0

- X and Y are not independent since

$$P(X = 2, Y = 2) = \frac{1}{6} \neq P(X = 2)P(Y = 2) = \frac{3}{16}$$

Joint Probability Density Function

A joint probability density function $f(x, y)$ for the continuous random variables X and Y for any region R of 2D space is

$$P((X, Y) \in R) = \int \int_R f(x, y) dxdy$$

- $f(x, y) \geq 0$ for all x,y and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dxdy = 1$
- Marginal pdfs

$$f(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

$$f(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

- The conditional probability density function of Y given X $=$ x is (for $f(x) > 0$)

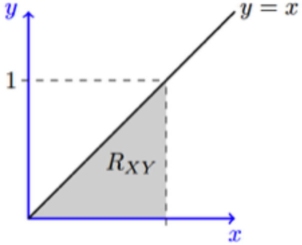
$$f_{Y|X}(y) = \frac{f(x, y)}{f(x)}$$

$$\int f_{Y|X}(y) dy = 1$$

Joint Probability Density Function. Example

Given that $f(x, y) = cx^2y$ for $0 \leq y \leq x \leq 1$ and 0 otherwise,

- the pdf is defined in the region R_{XY}



- constant $c = 1/10$ since

$$1 = \int_0^1 \int_0^x cx^2 y dy dx = \int_0^1 cx^4/2 dx = c/10$$

- the marginal pdf of Y is

$$f(y) = \int_y^1 10x^2 y dx = 10y(1 - y^3)/3$$