Chapter 5: JOINT PROBABILITY DISTRIBUTIONS

Part 1: Joint Discrete Probability Distributions...
Marginal Distributions
Conditional Distributions
Independence

Sections 5-1.1 to 5-1.4

Recall a discrete probability distribution (or probability mass function)

x	0	1	2
$f(x)$	0.50	0.20	0.30

Sometimes we're simultaneously interested in two or more discrete variables in a random experiment.

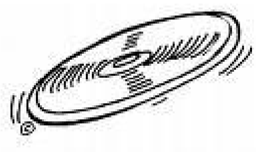
Examples

- Year in college vs. Number of credits taken
- Count of plants grown in a tray vs. Count of healthy plants
- Number of cigarettes smoked per day vs. Age of cancer onset

In general, if X and Y are two random variables, the probability distribution that defines their simultaneous behavior is called a joint probability distribution.

If X and Y are discrete, this distribution can be described with a joint probability mass function (this section).

If X and Y are continous, this distribution can be described with a joint probability density function (next section).

- **Example**: Plastic covers for CDs 

  Measurements for the length and width of a rectangular plastic covers for CDs are rounded to the nearest mm (so they are discrete).

  Let X denote the length and
  Y denote the width.

  The possible values of X are 129, 130, and 131 mm. The possible values of Y are 15 and 16 mm.

  Both X and Y are discrete.

There are 6 possible pairs (X, Y).

We show the probability for each pair in the following table:

		x= length		
		129	130	131
y=width	15	0.12	0.42	0.06
	16	0.08	0.28	0.04

The sum of all the probabilities is 1.0.

The combination with the highest probability is $(130, 15)$.

The combination with the lowest probability is $(131, 16)$.

The joint probability mass function is the function $f_{XY}(x, y) = P(X = x, Y = y)$. For example, we have $f_{XY}(129, 15) = 0.12$.

If we are given a joint probability distribution for X and Y, we can obtain the individual probability distribution for X or for Y...

- **Example**: Continuing plastic covers for CDs

  Find the probability that a CD cover has length of $129mm$ (i.e. X=129).

		x= length		
		129	130	131
y=width	15	**0.12**	0.42	0.06
	16	**0.08**	0.28	0.04

$$\begin{aligned} P(X=129) = & \; P(X=129 \text{ and } Y=15) \\ & + P(X=129 \text{ and } Y=16) \\ = & \; 0.12 \; + \; 0.08 = 0.20 \end{aligned}$$

  What is the probability distribution of X?

		x= length		
		129	130	131
y=width	15	0.12	0.42	0.06
	16	0.08	0.28	0.04
column totals		**0.20**	**0.70**	**0.10**

The probability distribution for X appears in the column totals...

x	129	130	131
$f_X(x)$	0.20	0.70	0.10

* NOTE: We've used a subscript X in the probability mass function of X, or $f_X(x)$, for clarification since we're considered more than one variable at a time now.

We can do the same for the Y random variable.

		x= length			**row totals**
		129	130	131	
y=width	15	0.12	0.42	0.06	**0.60**
	16	0.08	0.28	0.04	**0.40**
column totals		**0.20**	**0.70**	**0.10**	**1**

y	15	16
$f_Y(y)$	0.60	0.40

Because the the probability mass functions for X and Y appear in the margins of the table (i.e. column and row totals), they are often referred to as the **marginal distributions** for X and Y.

When there are two random variables of interest, we also use the term **bivariate probability distribution** or **bivariate distribution** to refer to the joint distribution.

- **Joint Probability Mass Function**

The joint probability mass function of the discrete random variables X and Y, denoted as $f_{XY}(x, y)$, satisfies

(1) $f_{XY}(x, y) \geq 0$

(2) $\sum_x \sum_y f_{XY}(x, y) = 1$

(3) $f_{XY}(x, y) = P(X = x, Y = y)$

- **Marginal Probability Mass Function**

If X and Y are discrete random variables with joint probability mass function $f_{XY}(x, y)$, then the marginal probability mass functions of X and Y are

$$f_X(x) = \sum_y f_{XY}(x, y)$$

and

$$f_Y(y) = \sum_x f_{XY}(x, y)$$

where the sum for $f_X(x)$ is over all points in the range of (X, Y) for which $X = x$ and the sum for $f_Y(y)$ is over all points in the range of (X, Y) for which $Y = y$.

When asked for $E(X)$ or $V(X)$ in a joint probability distribution problem, first calculate the marginal distribution $f_X(x)$ and work it as we did in chapter 3 for the univariate case (i.e. for a single random variable).

- **Example**: Batteries

Suppose that 2 batteries are randomly chosen without replacement from the following group of 12 batteries:

3 new
4 used (working)
5 defective

Let X denote the number of new batteries chosen.

Let Y denote the number of used batteries chosen.

a) Find $f_{XY}(x, y)$
{i.e. the joint probability distribution}.

ANS:
Though X can take on values 0, 1, and 2, and Y can take on values 0, 1, and 2, when we consider them jointly, $X + Y \leq 2$. So, not all combinations of (X, Y) are possible.

CASE: no new, no used (so all defective)

$$f_{XY}(0,0) = \frac{\binom{5}{2}}{\binom{12}{2}} = 10/66$$

CASE: no new, 1 used

$$f_{XY}(0,1) = \frac{\binom{4}{1}\binom{5}{1}}{\binom{12}{2}} = 20/66$$

CASE: no new, 2 used

$$f_{XY}(0,2) = \frac{\binom{4}{2}}{\binom{12}{2}} = 6/66$$

CASE: 1 new, no used

$$f_{XY}(1,0) = \frac{\binom{3}{1}\binom{5}{1}}{\binom{12}{2}} = 15/66$$

CASE: 2 new, no used

$$f_{XY}(2,0) = \frac{\binom{3}{2}}{\binom{12}{2}} = 3/66$$

CASE: 1 new, 1 used

$$f_{XY}(1,1) = \frac{\binom{3}{1}\binom{4}{1}}{\binom{12}{2}} = 12/66$$

x= number of *new* chosen

		0	1	2
y=number of	0	10/66	15/66	3/66
used	1	20/66	12/66	
chosen	2	6/66		

There are 6 possible (X,Y) pairs.

And, $\sum_x \sum_y f_{XY}(x,y) = 1$.

b) Find $E(X)$.

Conditional Probability Distributions

As we saw before, we can compute the conditional probability of an event *given* information of another event.

As stated before,

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

- **Example**: Continuing the plastic covers...

		x= length			row totals
		129	130	131	
y=width	15	0.12	0.42	0.06	0.60
	16	0.08	0.28	0.04	0.40
column totals		0.20	0.70	0.10	1

a) Find the probability that a CD cover has a length of $130mm$ GIVEN the width is $15mm$.

		x= length			row totals
		129	130	131	
y=width	15	0.12	0.42	0.06	0.60
	16	0.08	0.28	0.04	0.40
column totals		0.20	0.70	0.10	1

ANS: $P(X = 130|Y = 15) = \frac{P(X=130, Y=15)}{P(Y=15)}$

$$= 0.42/0.60 \ = \ 0.70$$

b) Find the conditional distribution of X given Y=15.

$$P(X = 129|Y = 15) = 0.12/0.60 \ = \ 0.20$$
$$P(X = 130|Y = 15) = 0.42/0.60 \ = \ 0.70$$
$$P(X = 131|Y = 15) = 0.06/0.60 \ = \ 0.10$$

Once you're GIVEN that Y=15, you're in a 'different space'.

We are now considering only the CD covers with a width of $15mm$. For this subset of the covers, how are the lengths (X) distributed.

The conditional distribution of X, or $f_{X|Y}(x)$, given Y=15:

x	129	130	131
$f_{X\|15}(x)$	0.20	0.70	0.10

Notice that the sum of these probabilities is 1, and this is a legitimate probability distribution .

* NOTE: Again, we use the subscript $X|Y$ for clarity to denote that this is a conditional distribution.

- **Conditional Probability Mass Function**

  Given discrete random variables X and Y with joint probability mass function $f_{XY}(x,y)$ the <u>conditional probability mass function</u> of Y given X=x is

$$f_{Y|X}(y) = \frac{f_{XY}(x,y)}{f_X(x)} \qquad \text{for} \qquad f_X(x) > 0.$$

The <u>conditional</u> probability is the *joint* probability over the *marginal* probability.

Notice that we can define $f_{X|Y}(x)$ in a similar manner if we are interested in that conditional distribution.

Because a conditional probability mass function $f_{Y|X}(y)$ is a probability mass function, the following properties are satisfied:

$$(1)\quad f_{Y|X}(y) \geq 0$$

$$(2)\quad \sum_y f_{Y|X}(y) = 1$$

$$(3)\quad f_{Y|X}(y) = P(Y = y|X = x)$$

- **Conditional Mean and Variance**

The conditional mean of Y given X=x, denoted as $E(Y|x)$ or $\mu_{Y|x}$ is

$$\begin{aligned} E(Y|x) &= \sum_y y f_{Y|X}(y) \\ &= \mu_{Y|x} \end{aligned}$$

and the conditional variance of Y given X=x, denoted as $V(Y|x)$ or $\sigma^2_{Y|x}$ is

$$\begin{aligned} V(Y|x) &= \sum_y (y - \mu_{Y|x})^2 f_{Y|X}(y) \\ &= \sum_y y^2 f_{Y|X}(y) - \mu^2_{Y|x} \\ &= E(Y^2|x) - [E(Y|x)]^2 \\ &= \sigma^2_{Y|x} \end{aligned}$$

- **Example**: Continuing the plastic covers...

		x= length			row totals
		129	130	131	
y=width	15	0.12	0.42	0.06	0.60
	16	0.08	0.28	0.04	0.40
column totals		0.20	0.70	0.10	1

a) Find the $E(Y|X = 129)$ and
$V(Y|X = 129)$.

ANS:
We need the conditional distribution first...

y	15	16
$f_{Y\|129}(y)$		

Independence

As we saw earlier, sometimes, knowledge of one event does not give us any information on the probability of another event.

Previously, we stated that if A and B were independent, then

$$P(A|B) = P(A).$$

In the framework of probability distributions, if X and Y are independent random variables, then $f_{Y|X}(y) = f_Y(y)$.

- **Independence**

  For discrete random variables X and Y, if any of the following properties is true, the others are also true, and X and Y are independent.

  (1) $f_{XY}(x,y) = f_X(x) f_Y(y)$ for all x and y

  (2) $f_{Y|X}(y) = f_Y(y)$
  for all x and y with $f_X(x) > 0$

  (3) $f_{X|Y}(x) = f_X(x)$
  for all x and y with $f_Y(y) > 0$

  (4) $P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$
  for any sets A and B in the range of X and Y.

Notice how (1) leads to (2):

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

- **Example**: Continuing the battery example

Two batteries were chosen without replacement.

Let X denote the number of new batteries chosen.

Let Y denote the number of used batteries chosen.

		x= number of *new* chosen		
		0	1	2
y=number	0	10/66	15/66	3/66
of used	1	20/66	12/66	
chosen	2	6/66		

a) Without doing any calculations, can you tell whether X and Y are independent?

- **Example**: Independent random variables

Consider the random variables X and Y, which both can take on values of 0 and 1.

		x		row totals
		0	1	
y	0	0.40	0.10	0.50
	1	0.40	0.10	0.50
column totals		0.80	0.20	1

a) Are X and Y independent?

y	0	1
$f_{Y\|0}(y)$		

y	0	1
$f_{Y\|1}(y)$		

Does $f_{Y|X}(y) = f_Y(y)$ for all x & y?

Does $f_{XY}(x, y) = f_X(x)f_Y(y)$ for all x & y?

		x		row totals
		0	1	
y	0	0.40	0.10	0.50
	1	0.40	0.10	0.50
column totals		0.80	0.20	1

i.e. Does $P(X = x, Y = y)$
$= P(X = x) \cdot P(Y = y)$?