Probability and Mathematical Statistics in Data Science

Lecture 13: 5.2: Functions of Random Variables Continued
Section 5.3: Methods of Indicators

Jointly Distributed Random Variables

- How can we model two random variables using probability models?

- We need to introduce joint probability distribution in order to model multiple random variables.

Joint PMF

- Let X and Y be two discrete random variables defined on the sample space. The joint probability mass function $p(x, y)$ is defined for each pair of numbers (x, y) by

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

- As in the single rv case, we must have $p(x, y) \geq 0$ and $\sum_x \sum_y p(x, y) = 1$

- Example:

p_{ij}	1	2	3
1	4/9	2/9	0
2	1/9	1/9	1/9

Marginal PMF

- The marginal probability mass functions of X and Y, denoted by $p_X(x)$ and $p_Y(y)$, respectively, are given by

$$p_X(x) = \sum_y p(x,y) \qquad p_Y(y) = \sum_x p(x,y)$$

- Example:

p_{ij}	1	2	3	$p_X(x)$
1	4/9	2/9	0	2/3
2	1/9	1/9	1/9	1/3
$p_Y(y)$	5/9	1/3	1/9	

- Notice that the marginal probability mass functions are automatically proper pmf's.

Joint Distributions : Example

- Let X = the deductible amount on the auto policy and Y = the deductible amount on the homeowner's policy. Possible (X, Y) pairs are then (100, 0), (100, 100), (100, 200), (250, 0), (250, 100), and (250, 200); the joint pmf specifies the probability associated with each one of these pairs, with any other pair having probability zero. Suppose the joint pmf is given in the accompanying **joint probability table**

$p(x, y)$		y = 0	100	200
x	100	.20	.10	.20
	250	.05	.15	.30

Marginal Distribution: Example

$p(x, y)$		y: 0	100	200
x	100	.20	.10	.20
	250	.05	.15	.30

- The possible X values are x =100 and x = 250, so computing row totals in the joint probability table yields:

- $p_X(100) = p(100, 0) + p(100, 100) + p(100, 200) = .50$

- $p_X(250) = p(250, 0) + p(250, 100) + p(250, 200) = .50$

Marginal Distribution: Example

The marginal pmf of X is then

$$p_X(x) = \begin{cases} .5 & x = 100, 250 \\ 0 & \text{otherwise} \end{cases}$$

Similarly, the marginal pmf of Y is obtained from column totals as

$$p_Y(y) = \begin{cases} .25 & y = 0, 100 \\ .50 & y = 200 \\ 0 & \text{otherwise} \end{cases}$$

Marginal Distribution

The **marginal probability mass function of X,** denoted by $p_X(x)$, is given by

$$p_X(x) = \sum_{y:\, p(x,y)>0} p(x, y) \quad \text{for each possible value } x$$

Similarly, the **marginal probability mass function of Y** is

$$p_Y(y) = \sum_{x:\, p(x,y)>0} p(x, y) \quad \text{for each possible value } y.$$

Two random variables X and Y are said to be **independent** if for every pair of x and y values

$$p(x, y) = p_X(x) \cdot p_Y(y) \qquad \text{when } X \text{ and } Y \text{ are discrete}$$

Question

A service station has both self-service and full-service islands. On each island, there is a single regular unleaded pump with two hoses. Let X denote the number of hoses being used on the self-service island at a particular time, and let Y denote the number of hoses on the full-service island in use at that time. The joint pmf of X and Y appears in the accompanying tabulation.

$p(x, y)$		y = 0	1	2
	0	.10	.04	.02
x	1	.08	.20	.06
	2	.06	.14	.30

a. What is $P(X = 1 \text{ and } Y = 1)$?
b. Compute $P(X \leq 1 \text{ and } Y \leq 1)$.
c. Give a word description of the event $\{X \neq 0 \text{ and } Y \neq 0\}$, and compute the probability of this event.
d. Compute the marginal pmf of X and of Y. Using $p_X(x)$, what is $P(X \leq 1)$?
e. Are X and Y independent rv's? Explain.

Additivity of Expectation

- This is a very useful property – no matter what the joint distribution of X and Y may be, we have:

$$E(X + Y) = E(X) + E(Y)$$

- Whether X and Y are dependent or independent, this holds, making it enormously useful.

- We also have linearity: $E(aX + bY) = aE(X) + bE(Y)$

Method of indicators

- Recall that we talked about "classifying and counting" – so, we divide up the outcomes into those that we are interested in (successes), and everything else (failures), and then count the number of successes.
- We can represent these outcomes as 0 and 1, where 1 marks a success and 0 and failure, so if we model the trials as draws from a box, we can count the number of successes by counting up the number of times we drew a 1.
- We can represent each draw as a Bernoulli trial, where $p = P(S)$

Using indicators and additivity

- For example, say we roll a die 10 times, and success is rolling a 1.

- Then p=1/6, and we can define a Bernoulli rv as $X = \begin{cases} 0, & w.p. 5/6 \\ 1, & w.p. 1/6 \end{cases}$

- We can also define an event A: let A be the event of rolling a 1 and define a random variable I_A (indicator of an event) that takes the value 1 if A occurs and 0 otherwise.

- This is a Bernoulli random variable, what is its expectation?

- Now let $X \sim Bin(10, \frac{1}{6})$, so X counts the number of successes in 10 rolls. Let's find $E(X)$ using additivity and indicators:

Using indicators

Exercise 5.7.6: A die is rolled 12 times. Find the expectation of:

a) the number of times the face with five spots appears

b) the number of times an odd number of spots appears

c) the number of faces that don't appear

d) the number of faces that do appear