Joint discrete random variables

November 16, 2020

- Probabilistic models often involve several random variables of interest.
- in a medical diagnosis context, the results of several tests may be significant
- in a networking context, the workloads of several gateways may be of interest.

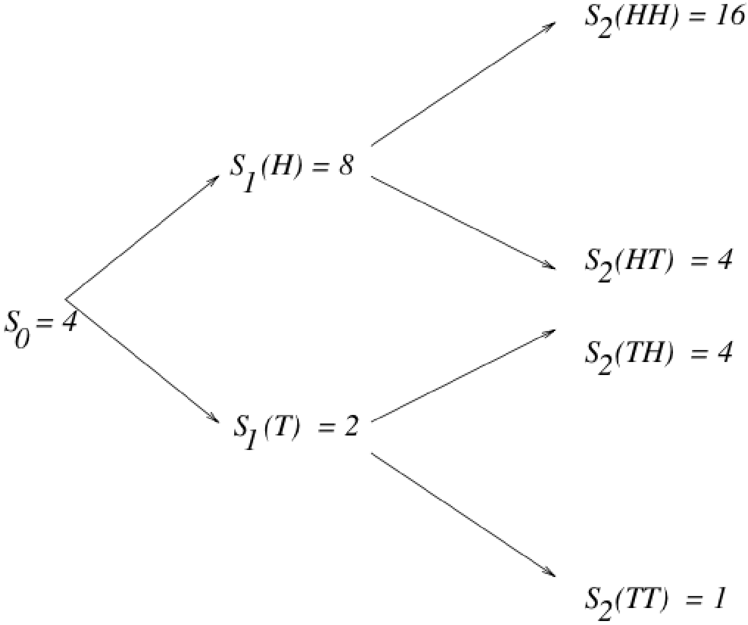
- their values may relate in interesting ways.
- consider probabilities involving simultaneously the numerical values of several random variables and investigate their mutual couplings.

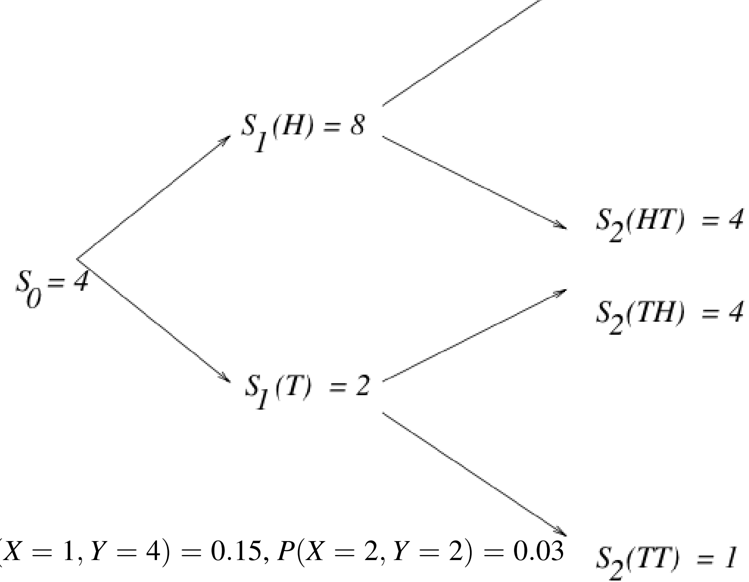
The joint p.m.f of the discrete random variables X and Y is defined by

$$p_{X,Y}(x, y) = P(X = x, Y = y)$$

1. $p_{X,Y}(x, y) \geq 0$
2. $\sum_{x,y} p(x, y) = 1$

Calls are made to check the airline schedule at your departure city. You monitor the number of bars of signal strength on your cell phone and the number of times you have to state the name of your departure city before the voice system recognizes the name


$S_2(HH) = 16$
$S_1(H) = 8$
$S_2(HT) = 4$
$S_0 = 4$
$S_2(TH) = 4$
$S_1(T) = 2$
$S_2(TT) = 1$



$P(X = 1, Y = 4) = 0.15, P(X = 2, Y = 2) = 0.03$

Binomial tree of stock prices with $S_0 = 4, u = 1/d = 2.$

Example

Suppose that 3 batteries are randomly chosen from a group of 3 new, 4 used but still working, and 5 defective batteries. If we let X and Y denote, respectively, the number of new and used but still working batteries that are chosen, find the joint probability mass function of X and Y

Solution

- $X \in \{0, 1, 2, 3\}$ and $Y \in \{0, 1, 2, 3\}$
- $$P(X = 0, Y = 0) = P(3 \text{ defective}) = \frac{\binom{5}{3}}{\binom{12}{3}}$$

- $X \in \{0, 1, 2, 3\}$ and $Y \in \{0, 1, 2, 3\}$
- $$P(X = 0, Y = 0) = P(3 \text{ defective}) = \frac{\binom{5}{3}}{\binom{12}{3}}$$

$$
\begin{aligned}
&P(X=0, Y=1) \\
&= P(0 \text{ new}, 1 \text{ used }, 2 \text{ defective}) \\
&= \frac{\binom{4}{1}\binom{5}{2}}{\binom{12}{3}}
\end{aligned}
$$

Solution

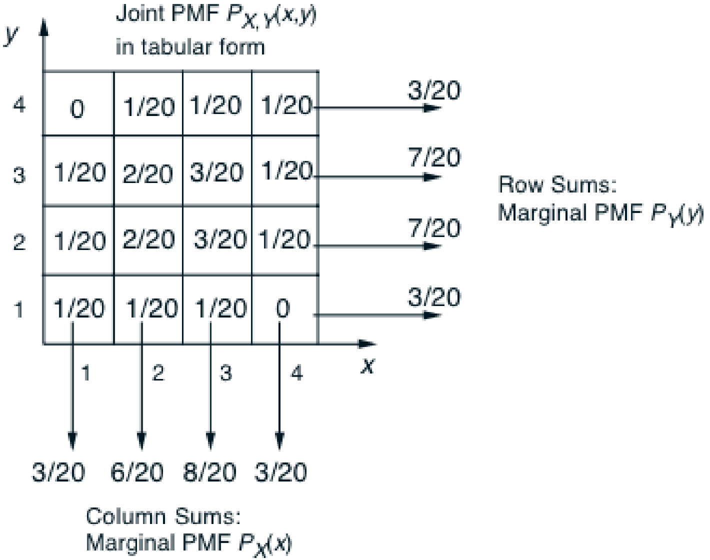
TABLE 4.1 $P\{X = i, Y = j\}$

$i \backslash j$	0	1	2	3	Row Sum $= P\{X = i\}$
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$	$\frac{84}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0	$\frac{108}{220}$
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0	$\frac{27}{220}$
3	$\frac{1}{220}$	0	0	0	$\frac{1}{220}$
Column Sums $= P\{Y = j\}$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	

By total probability formula

$$p_X(x) = \sum_y P(X = x, Y = y) = \sum_y p_{X,Y}(x, y)$$

$$p_Y(y) = \sum_x p_{X,Y}(x, y)$$


Joint PMF $P_{X,Y}(x,y)$
in tabular form
y
4
0
1/20
1/20
1/20
3
1/20
2/20
3/20
1/20
2
1/20
2/20
3/20
1/20
1
1/20
1/20
1/20
0
3/20
7/20
7/20
3/20
Row Sums:
Marginal PMF $P_Y(y)$
1
2
3
4
x
3/20
6/20
8/20
3/20
Column Sums:
Marginal PMF $P_X(x)$

1. Joint cdf

$$F_{X,Y}(x,y) = P(X \le x, Y \le y)$$

2. Marginal cdf
   - $P(X \le x) = F_X(x)$
     $= F_{X,Y}(X \le x, Y \le \infty)$
   - $P(Y \le y) = F_Y(y)$
     $= F_{X,Y}(X \le \infty, Y \le y)$

1. Joint cdf

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$$

2. Marginal cdf
   - $P(X \leq x) = F_X(x)$
     $= F_{X,Y}(X \leq x, Y \leq \infty)$
   - $P(Y \leq y) = F_Y(y)$
     $= F_{X,Y}(X \leq \infty, Y \leq y)$

- $$E(g(X,Y)) = \sum_{x,y} g(x,y)p(x,y)$$

- Sum of RV

$$E(X_1 + \ldots + X_n) = E(X_1) + \ldots + E(X_n)$$

Toss a fair 6-sided dice twice. Let X and Y be the number of the first and second roll. Then expected value of the sum of two numbers $E(X+Y) = E(X) + E(Y) = 3.5 + 3.5 = 7$

Your probability class has 300 students and each student has probability 1/3 of getting an A, independently of any other student. What is the mean of X, the number of students that get an A?

Solution

- $X_i = \begin{cases} 1 \text{ if the } i\text{th student gets A} \\ 0 \text{ otherwise} \end{cases}$
- $X_i \sim Bino(1/3), E(X_i) = 1/3$
- $X = X_1 + ... + X_{100}$
- $E(X) = E(X_1) + ... + E(X_{100}) = 100.\frac{1}{3} = \frac{100}{3}$

Suppose that n people throw their hats in a box and then each picks up one hat at random. What is the expected value of X, the number of people that get back their own hat?

Solution

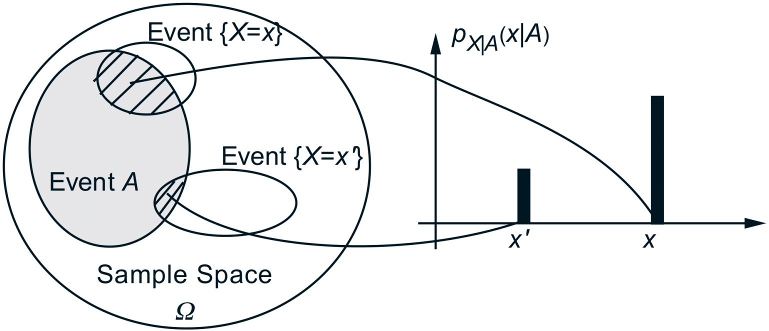
- $X_i = \begin{cases} 1 \text{ if the person } i \text{ takes his/her own hat} \\ 0 \text{ otherwise} \end{cases}$
- $P(X_i = 1) = 1/n, P(X_i = 0) = 1 - 1/n,\ E(X_i) = 1/n$
- $X = X_1 + ... + X_n$
- $E(X) = E(X_1) + ... + E(X_n) = n.\frac{1}{n} = 1$

Conditioning

The conditional pmf of a RV X given an event A is

$$p_{X|A}(x) = P(X = x|A) = \frac{P((X = x) \cap A)}{P(A)}$$

if $P(A) > 0$



Event {X=x}
Event {X=x'}
Event A
Sample Space
Ω
p_{X|A}(x|A)
x'
x

$$\sum_x p_{X|A}(x) = 1$$

Let X be the roll of a fair die and let A be the event that the roll is an even number. Then

$$p_{X|A}(1) = \frac{P(X = 1 \text{ and roll is even})}{P(\text{roll is even})} = 0$$

Let X be the roll of a fair die and let A be the event that the roll is an even number. Then

$$p_{X|A}(2) = \frac{P(X = 2 \text{ and roll is even})}{P(\text{roll is even})} = \frac{1}{3}$$

$$p_{X|A}(i) = \begin{cases} 0, & i = 1, 3, 5 \\ \frac{1}{3}, & 2, 4, 6 \end{cases}$$

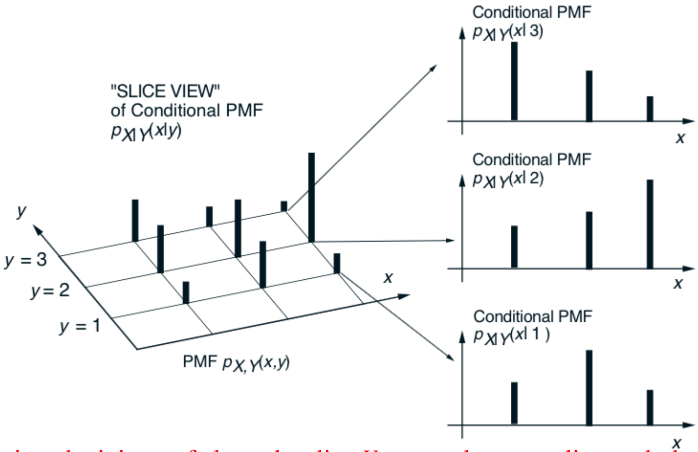
Let X be the roll of a fair die and let A be the event that the roll is an even number. Then

$$p_{X|A}(2) = \frac{P(X = 2 \text{ and roll is even})}{P(\text{roll is even})} = \frac{1}{3}$$

$$p_{X|A}(i) = \begin{cases} 0, & i = 1, 3, 5 \\ \frac{1}{3}, & 2, 4, 6 \end{cases}$$

- 2 RVs X and Y
- given $Y = y$ with $P(Y = y) > 0$
- conditional pmf of X

$$
\begin{aligned}
p_{X|Y}(x|y) &= P(X = x|Y = y) \\
&= \frac{P(X = x, Y = y)}{P(Y = y)}
\end{aligned}
$$



For each y, we view the joint pmf along the slice $Y = y$ and renormalize such that $\sum_x p_{X|Y}(x|y) = 1$

Example

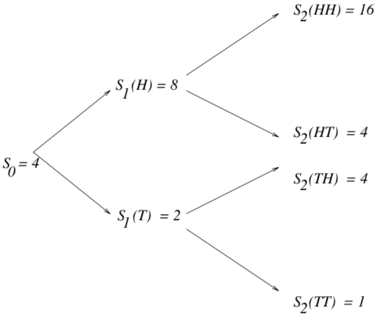
	x = number of bars of signal strength			
y = number of times city name is stated	**1**	**2**	**3**	**Marginal probability distribution of Y**
4	0.15	0.1	0.05	0.3
3	0.02	0.1	0.05	0.17
2	0.02	0.03	0.2	0.25
1	0.01	0.02	0.25	0.28
	0.2	0.25	0.55	
	Marginal probability distribution of X			

$$P(X = 2|Y = 1) = \frac{P(X=2,Y=1)}{P(Y=1)} = \frac{0.02}{0.01+0.02+0.25} = \frac{2}{28}$$

u: up factor, d: down factor

- Initial stock price S_0
- Next period
  - **Upward**: uS_0
  - **Downward**: dS_0

  where $0 < d < 1 < u$, $(d = \frac{1}{u})$
- Toss a coin
  - Head: move up with prob $p = 1/3$
  - Tail: move down with prob $1 - p$
- Find conditional pmf of S_3 given that $S_2 = 4$



Binomial tree of stock prices with $S_0 = 4$, $u = 1/d = 2$.

Two RVs X and Y are **independent** if

$$p_{X,Y}(x,y) = p_X(x)p_Y(y)$$

for all x, y.
Equivalent condition

$$F_{X,Y}(x,y) = F_X(x)F_Y(y), \forall x,y$$

Equivalent condition for independence

$$p_{X|Y}(x|y) = p_X(x)$$

or

$$p_{Y|X}(y|x) = p_Y(y)$$

for all x, y

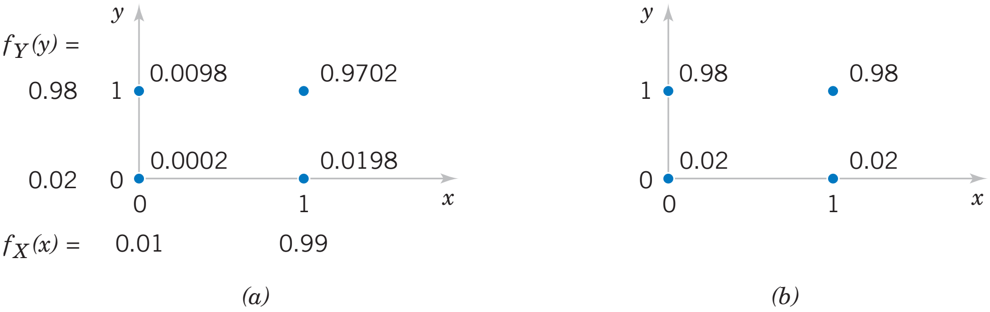
knowledge of the values of X does not change any of the probabilities associated with the values for Y and inversely.

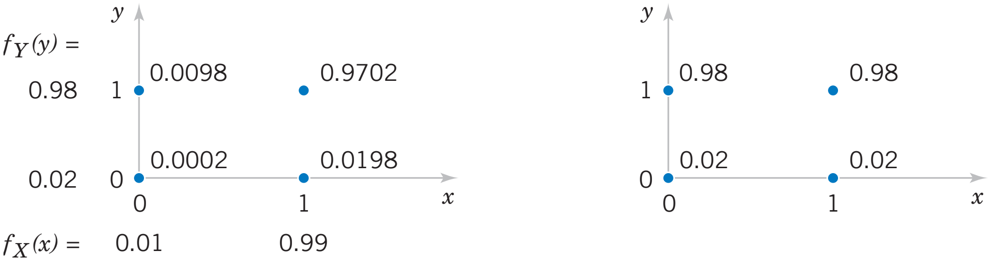
In a plastic molding operation, each part is classified as to whether it conforms to color and length specifications.

$$X = \begin{cases} 1 & \text{if the part conforms to color specification} \\ 0 & \text{otherwise} \end{cases}$$

$$Y = \begin{cases} 1 & \text{if the part conforms to length specification} \\ 0 & \text{otherwise} \end{cases}$$

(a) joint pmf and (b) conditional pmf of Y given $X = x$

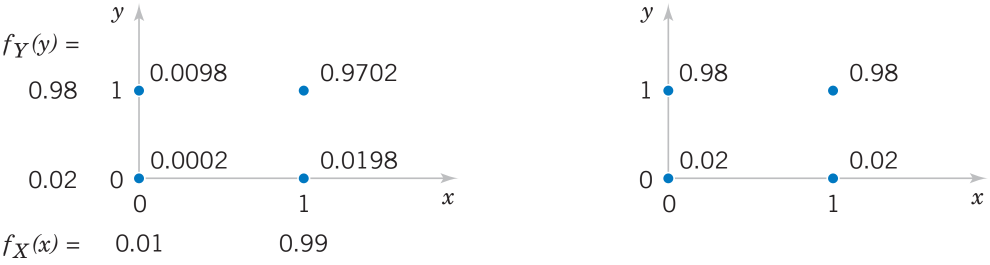




(a) (b)

Given $X = 1$, we have

$$p_{Y|X}(0|1) = p_Y(0) = .98, p_{Y|X}(1|1) = p_Y(1) = .02$$



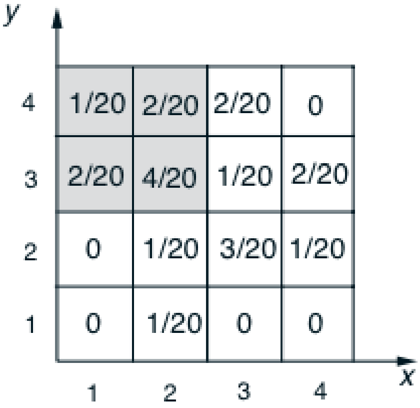
(a) (b)

Given $X = 0$, we have

$$p_{Y|X}(0|0) = p_Y(0) = .98, p_{Y|X}(1|0) = p_Y(1) = .02$$

$$p_{Y|X}(y|x) = p_Y(y)\ \forall x, y$$

Hence X and Y are independence



X and Y are not independent because

$$P_{X|Y}(1|1) = 0 \neq P_X(1)$$

If two RVs X and Y are independent then

- $E(XY) = E(X)E(Y)$
- $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$
- $f(X)$ and $g(Y)$ are independent for any functions f, g

$X_1, X_2, ..., X_n$ are independent

- $E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$
- $\text{Var}(X_1 + X_2 + \ldots + X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \ldots + \text{Var}(X_n)$

Example - Mean and Variance of Binomial

If $X \sim Bino(n, p)$ then

$$X = X_1 + X_2 + \ldots + X_n$$

where X_i are independently and identially distributed (i.i.d) $Ber(p)$

$$E(X) = E(X_1) \ldots + E(X_n) = nE(X_1) = np$$

$$\mathrm{Var}(X) = n\,\mathrm{Var}(X_1) = np(1-p)$$

If $X \sim Bino(n,p)$ then

$$X = X_1 + X_2 + \ldots + X_n$$

where X_i are independently and identially distributed (i.i.d) $Ber(p)$

$$E(X) = E(X_1)... + E(X_n) = nE(X_1) = np$$
$$\mathrm{Var}(X) = n\,\mathrm{Var}(X_1) = np(1-p)$$