Multiple Discrete Random Variables

Random Variable

Experiment (Physical Model)

→ Compose of procedure & observation

→ From observation, we get outcomes

→ From all outcomes, we get a (mathematical) probability model called "Sample space"

→ From the model, we get $P[A]$, $A \subset S$

Random Variable

From a probability model

- Ex.: 2 traffic lights, observe the seq. of light

  $S = \{R_1R_2, R_1G_2, G_1R_2, G_1G_2\}$
- If assign a number to each outcome in S, each number that we observe is called "**Random Variable**"
- Observe the number of red light

  $S_X = \{0,1,2\}$

How about Observe more than one thing in an experiment ?

What is Multiple Discrete RV?

- Each observation → Random Variable
- 2 observations → 2 Random Variables
- ≥ 2 observations → Multiple RVs

Joint Probability Mass Function

- For an experiment, Observe one thing
  - Model with one Random Variable
  - Describe the prob. model by using PMF
- For the same experiment, Observe 2 things
  - 2 Random Variables → X and Y
  - Joint PMF
- $P_{X,Y}(x,y)$

Probability Mass Function

$$P_X(x) = P[X=x]$$

PMF of random variable X

$X(s) = x$

Event of all outcomes s in S

A Random variable **X**

X is a dummy variable

$P_X(a)$ or $P_X(z)$ or $P_X(☺)$

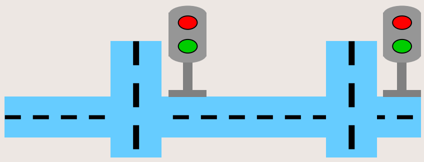
Joint PMF

Definition:

$$P_{X,Y}(x,y) = P[X=x,\ Y=y]$$

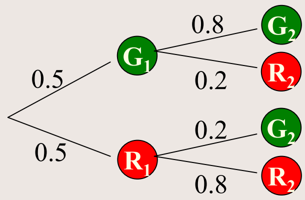
$$S_{X,Y} = \{(x,y) \mid P_{X,Y}(x,y) > 0\}$$

Example



- Timing coordination of 2 traffic lights
  - P[the second light is the same color as the first when the first light is given] = 0.8
  - Assume 1^{st} light is equally likely to be green or red

Example



$P[G_1G_2] = 0.4$

$P[G_1R_2] = 0.1$

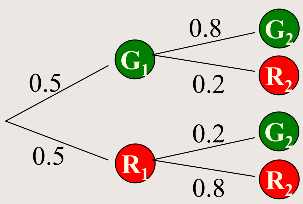
$P[R_1G_2] = 0.1$

$P[R_1R_2] = 0.4$

- Find P[The second light is green] ?
- Find P[wait for at least one light] ?
- Let observe
  - **number of G** and **number of G before 1st R**
  - Find the **Joint PMF**

Example

- Let
  - Count number of G $\rightarrow$ random variable X
  - Count number of G before 1st R $\rightarrow$ Y



$P[G_1G_2] = 0.4$ $X = 2, Y = 2$

$P[G_1R_2] = 0.1$ $X = 1, Y = 1$

$P[R_1G_2] = 0.1$ $X = 1, Y = 0$

$P[R_1R_2] = 0.4$ $X = 0, Y = 0$

Example

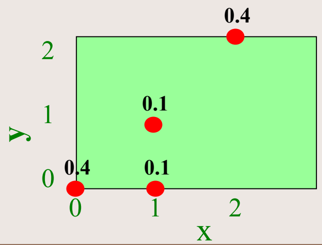
- Let g(s) transforms each outcome → a pair of RV (X,Y)
  - $g(G_1G_2) = (2,2)$ $g(G_1R_2) = (1,1)$
  - $g(R_1G_2) = (1,0)$ $g(R_1R_2) = (0,0)$
- For each pair of x,y
  - $P_{X,Y}(x,y)$ = sum of prob. that X = x and Y = y
  - $P_{X,Y}(1,0) \rightarrow P[R_1G_2]$

Example

- Joint PMF can be written in 3 forms:

$$P_{X,Y}(x,y) = \begin{cases} 0.4 & x=2, y=2 \\ 0.1 & x=1, y=1 \\ 0.1 & x=1, y=0 \\ 0.4 & x=0, y=0 \\ 0 & \text{Otherwise} \end{cases}$$

$P_{X,Y}(x,y)$	y=0	y=1	y=2
x=0	0.4	0	0
x=1	0.1	0.1	0
x=2	0	0	0.4



Joint PMF properties

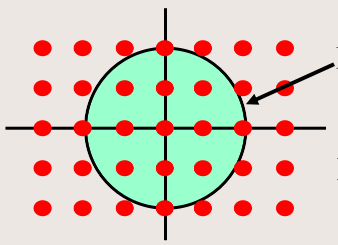
$$\sum_{x \in S_X} \sum_{y \in S_Y} P_{X,Y}(x,y) = 1$$

$$P_{X,Y}(x,y) \geq 0$$

Theorem

For any subset $B \subset S$ of X-Y plane,
the probability of the event $\{(X,Y) \in B\}$ is

$$P[B] = \sum_{(x,y) \in B} P_{X,Y}(x,y)$$

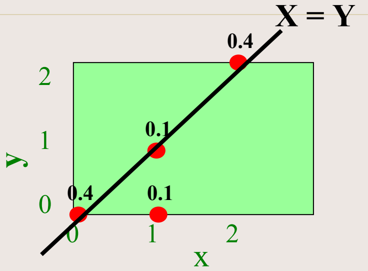


$B = \{X^2 + Y^2 \leq 2^2\}$

Subset B of (X,Y) plane
Point $(X,Y) \in S_{X,Y}$

Example

$P_{X,Y}(x,y)$	y=0	y=1	y=2
x=0	0.4	0	0
x=1	0.1	0.1	0
x=2	0	0	0.4



B = event that X equals Y

$P[B] = P_{X,Y}(0,0) + P_{X,Y}(1,1) + P_{X,Y}(2,2)$
$= 0.4 + 0.1 + 0.4 \quad = \; 0.9$

C = event that X > Y

$P[C] = P_{X,Y}(1,0) \; = 0.1$

Marginal PMF

- In an experiment with 2 RVs, X and Y
  - Possible to consider only one (X) and ignore Y
  - $P_X(x)$

Theorem: For random variables X and Y with joint PMF $P_{X,Y}(x,y)$:

$$\mathbf{P_X(x) = \sum_y P_{X,Y}(x,y)}$$

$$\mathbf{P_Y(y) = \sum_x P_{X,Y}(x,y)}$$

Marginal PMF Example

$P_{X,Y}(x,y)$	y=0	y=1	y=2
x=0	0.4	0	0
x=1	0.1	0.1	0
x=2	0	0	0.4

- Find Marginal PMF of X and Y
- $S_X = \{0,1,2\}$ $S_Y = \{0,1,2\}$

$$\mathbf{P_X(0)} = \sum_{y=0}^{2} P_{X,Y}(0,y) = 0.4 \qquad \mathbf{P_X(2)} = \sum_{y=0}^{2} P_{X,Y}(2,y) = 0.4$$

$$\mathbf{P_X(1)} = \sum_{y=0}^{2} P_{X,Y}(1,y) = 0.1+0.1 \qquad \mathbf{P_X(x)} = 0 \quad x \neq 0,1,2$$

Marginal PMF

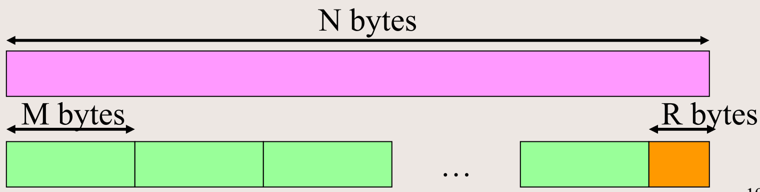
Margin

$P_{X,Y}(x,y)$	y=0	y=1	y=2	$P_X(x)$
x=0	0.4	0	0	0.4
x=1	0.1	0.1	0	0.2
x=2	0	0	0.4	0.4
$P_Y(y)$	0.5	0.1	0.4	1

Margin

Example (Leon-Garcia)

- The number of bytes, N, in a message is geometric distribution with parameter (1-p)
- A maximum packet size = M bytes
- Let Q = the number of packets
- Let R = the number of left over bytes



Example

- Find Joint Probability Mass Function $P_{Q,R}(q,r)$

Solution:

- N, Q, R, and M
- N = ?

  $N = QM + R$
- S_N, S_Q, S_R = ?

  $S_N = \{0,1,2,3,\ldots\}$

  $S_Q = \{0,1,2,3,\ldots\}$

  $S_R = \{0,1,2,3,\ldots,M-1\}$

Example

- $P_{Q,R}(q,r) = P_N(n)$

  $= P[N = n]$

  $= P[N = qM + r]$
- N = Geometric RV with parameter $(1 – p)$

Note: Geometric RV

- Two version of Geometric RV

  $S_X = \{1,2,3,\ldots\}$

  $S_Y = \{0,1,2,\ldots\}$
- E[X] = E[Y] ???

  → No

  $E[X] = 1/p$

  $E[Y] = (1-p)/p$

Geometric Random Variable

Definition: X is a Geometric Random Variable if the PMF of X, $P_X(x)$, has the form:

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & x = 1,2,3,\ldots \\ 0 & \text{Otherwise} \end{cases}$$

where p = (0,1)

$$P_Y(y) = \begin{cases} p(1-p)^{y} & y = 0,1,2,\ldots \\ 0 & \text{Otherwise} \end{cases}$$

Example

- $P_{Q,R}(q,r) = P_N(n)$

  $= P[N = n]$

  $= P[N = qM + r]$
- $P_N(n) = (1 - p)\,(1 - (1 - p))^n$

  $= (1 - p)\,(1 - (1 - p))^{qM + r}$

  $= (1 - p)\,p^{qM + r}$
- $P_{Q,R}(q,r) = (1 - p)\,p^{qM + r}$

Example

- Find $P_Q(q) = ?$
- $P_Q(q) = \sum_{r=0}^{M-1} (1-p)\, p^{qM+r}$

$$= (1-p)\, p^{qM} \sum_{r=0}^{M-1} p^r$$

$$= (1-p)\, p^{qM} \frac{1-p^M}{1-p}$$

$$= (1-p^M)\,(p^M)^q \qquad q = 0,1,2,3,\ldots$$

Example

- Find $P_R(r) = ?$
- $P_R(r) = \sum_{q=0}^{\infty} (1-p)p^{qM+r}$

$$= (1-p)p^r \sum_{q=0}^{\infty} p^{qM}$$

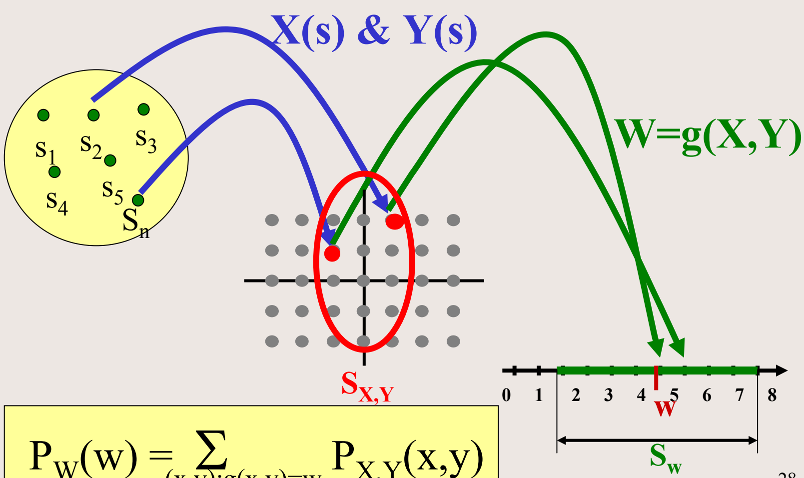
$$= (1-p)p^r \frac{1}{1-p^M}$$

$$= \frac{(1-p)}{(1-p^M)}\, p^r \qquad r = 0,1,2,\ldots,M-1$$

Marginal PMF

- From $P_{X,Y}(x,y)$, we can find
  - $P_X(x)$
  - $P_Y(y)$
- From $P_X(x)$ or $P_Y(y)$, can we find $P_{X,Y}(x,y)$?
  - NO

Derived Random Variable
Functions of 2 RVs



$$P_W(w) = \sum_{(x,y):g(x,y)=w} P_{X,Y}(x,y)$$

Example

- Fax Sending (text-40sec & graphics-60sec)

$P_{S,T}(s,t)$	t = 40	t = 60
s = 1 sheet	0.15	0.1
s = 2 sheets	0.3	0.2
s = 3 sheets	0.15	0.1

Let D = duration for sending one fax

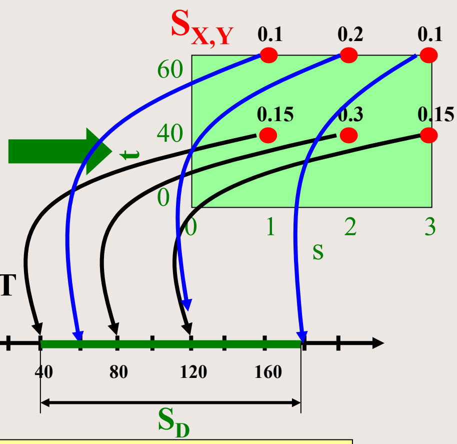
= g(S,T)

= ST

Find $P_D(d)$, S_D, and E[D]

Example

$P_{S,T}(s,t)$	t = 40	t = 60
s = 1 sheet	0.15	0.1
s = 2 sheets	0.3	0.2
s = 3 sheets	0.15	0.1



D = g(S,T) = ST

S_D = {40,60,80,120,180}

Example

$S_D=\{40,60,80,120,180\}$

$$P_D(d) = \sum_{(s,t):g(s,t)=d} P_{S,T}(s,t)$$

$$P_D(d) = \begin{cases} 0.15 & d = 40 \\ 0.1 & d = 60 \\ 0.3 & d = 80 \\ 0.15 + 0.2 & d = 120 \\ 0.1 & d = 180 \\ 0 & \text{Otherwise} \end{cases}$$

$$E[D] = \sum_{d \in S_D} dP_D(d)$$

$$= \ 96 \text{ sec}$$

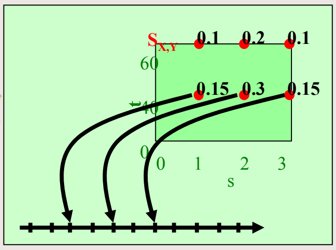
Expected Value of g(X,Y)

Theorem: for W = g(X,Y)

$$E[W] = \sum_{x \in S_X} \sum_{y \in S_Y} g(X,Y) \ P_{X,Y}(x,y)$$

From Last Example

To find the E[D], D = g(S,T) = ST



$$P_D(d) = \begin{cases} 0.15 & d = 40 \\ 0.1 & d = 60 \\ 0.3 & d = 80 \\ 0.15 + 0.2 & d = 120 \\ 0.1 & d = 180 \\ 0 & \text{Otherwise} \end{cases}$$

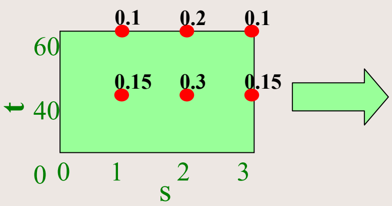
$$E[D] = \sum_{d \in S_D} dP_D(d)$$

Map g(S,T) → D   Find $P_D(d)$ Find E[D]

From Last Example

With the theorem, we can directly find E[D]

$$E[D] = \sum_{s=1}^{3} \sum_{t=40,60} st\ P_{S,T}(s,t)$$



E[D] = 1*40*0.15 + 1*60*0.1 +
2*40*0.3 + 2*60*0.2 +
3*40*0.15 + 3*60*0.1
= 96 sec

Expectations

- $E[W]$ for $W = g(X,Y)$
- $E[X+Y]$
- $Var[X+Y]$ (Variance of sum of 2 RVs)
- $Cov[X,Y]$ (Covariance)
- $r_{X,Y}$ (Correlation)
- $\rho_{X,Y}$ (Correlation Coefficient)

For any 2 RVs

Theorem:

$$\boxed{E[X + Y] = E[X] + E[Y]}$$

- Find E[X] and E[Y]
  → Marginal PMF

Var[X+Y]

Definition: $Var[X] = E[(X - \mu_x)^2]$

$$
\begin{aligned}
Var[X+Y] &= E[((X+Y) - \mu_{X+Y})^2] \\
&= E[((X+Y) - (\mu_X + \mu_Y))^2] \\
&= E[((X-\mu_x) + (Y-\mu_Y))^2] \\
&= E[(X-\mu_x)^2 + 2(X-\mu_X)(Y-\mu_Y) + (Y-\mu_Y)^2] \\
&= E[(X-\mu_x)^2] + 2E[(X-\mu_X)(Y-\mu_Y)] + E(Y-\mu_Y)^2]
\end{aligned}
$$

Theorem:

$Var[X+Y] = Var[X] + Var[Y] + 2\,E[(X-\mu_x)(Y-\mu_Y)]$

Covariance

Covariance of X and Y

Definition: $Cov[X,Y] = E[(X-\mu_x)(Y-\mu_Y)]$

$$
\begin{aligned}
Cov[X,Y] &= E[XY - \mu_x Y - \mu_Y X + \mu_x \mu_y] \\
&= E[XY] - \mu_x E[Y] - \mu_Y E[X] + \mu_x \mu_y \\
&= E[XY] - \mu_x \mu_y - \mu_x \mu_y + \mu_x \mu_y
\end{aligned}
$$

Theorem: $Cov[X,Y] = E[XY] - \mu_x \mu_y$

Covariance of X and Y

Theorem: $Cov[X,Y] = E[XY] - \mu_x \mu_y$

Correlation

If $X = Y \Rightarrow Cov[X,X] = E[XX] - \mu_x \mu_x$

$= E[X^2] - \mu_x^{\ 2}$

$= E[X^2 - 2\,\mu_x^{\ 2} + \mu_x^{\ 2}]$

$= E[X^2 - 2\,\mu_x X + \mu_x^{\ 2}]$

$= E[(X - \mu_x)^2]$

$= Var[X]$

If μ_x or $\mu_Y = 0 \Rightarrow Cov[X,Y] = E[XY]$

Correlation

Definition: The correlation of X and Y is $r_{X,Y}$

$r_{X,Y} = E[XY]$

Theorem: $Cov[X,Y] = r_{X,Y} - \mu_x \mu_y$

More Definition

Definition 1:

X and Y are **Orthogonal** if $r_{X,Y} = 0$; $E[XY]=0$

Definition 2:

X and Y are **Uncorrelated** if $Cov[X,Y] = 0$

Definition 3:

Correlation Coefficient of X and Y is

$$\rho_{X,Y} = \frac{Cov[X,Y]}{\sqrt{Var[X]Var[Y]}} = [-1, 1]$$

Correlation Coefficient

- $\rho_{X,Y}$
  - Describes the info about Y by observing X
- $\rho_{X,Y} > 0$
  - If X $\Uparrow$ (relative to mean) → Y $\Uparrow$
  - If X $\Downarrow$ (relative to mean) → Y $\Downarrow$
- $\rho_{X,Y} < 0$
  - If X $\Uparrow$ (relative to mean) → Y $\Downarrow$
  - If X $\Downarrow$ (relative to mean) → Y $\Uparrow$
- Example:
  - X = student's height, Y= student's weight $\rho_{X,Y} > 0$
  - X = cell phone distance, Y= Rx signal Strength $\rho_{X,Y} < 0$

Uncorrelated

If X and Y are Independent, then

➔ $Cov[X,Y] = 0$ ➔ $\rho_{X,Y} = 0$

➔ X and Y are Uncorrelated

Note:

If X and Y are Uncorrelated,

➔ X and Y **may or may not** Independent

Example

$P_{S,T}(s,t)$	t = 40	t = 60
s = 1 sheet	0.15	0.1
s = 2 sheets	0.3	0.2
s = 3 sheets	0.15	0.1

Quiz 4.7 Find

(1) E[S] and Var[S]

(2) E[T] and Var[T]

(3) $r_{S,T} = E[ST]$

(4) Cov[S,T]

(5) $\rho_{S,T}$

Conditional Joint PMF by an Event

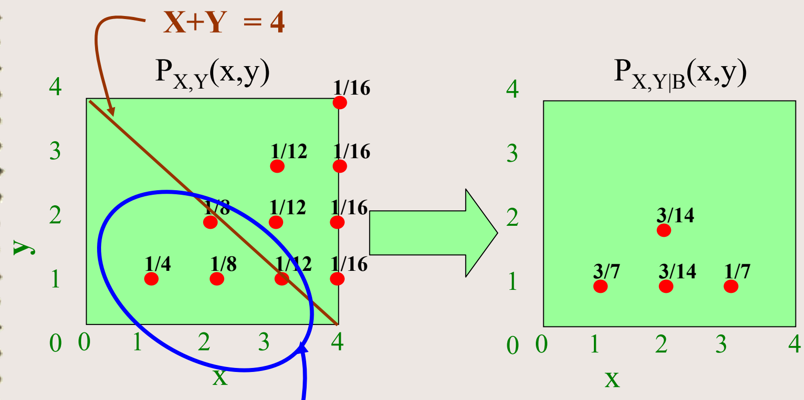
$$P_{X,Y|B}(x,y) = \frac{P[(X=x, Y=y) \cap B]}{P[B]}$$

If $(X=x, Y=y) \in B$ $\rightarrow$ $(X=x, Y=y) \cap B = (X=x, Y=y)$

$$P_{X,Y|B}(x,y) = \begin{cases} \dfrac{P[(X=x, Y=y)]}{P[B]} & (x,y) \in B \\ 0 & \text{Otherwise} \end{cases}$$

Example $P_{X,Y|B}(x,y)$

Let $B = \{X+Y \leq 4\}$ Find $\mathbf{P_{X,Y|B}(x,y)}$



$B = \{X+Y \leq 4\}$ $\rightarrow$ $P[B] = 7/12$

Conditional PMF

- Special case of Conditional Joint PMF by an Event
  → the Event is X=x or Y=y
- $P_{X,Y|B}(x,y)$ when $B = \{Y=y\}$
  → $P_{X,Y|Y=y}(x,y) = P_{X|Y}(x|y)$

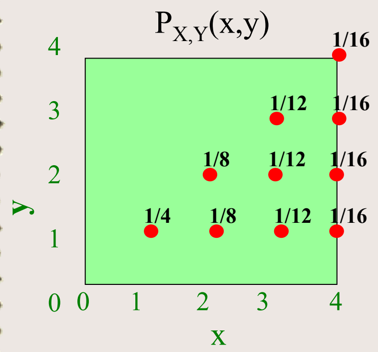
Definition: $P_{X|Y}(x|y) = P[X=x \mid Y=y]$

Conditional PMF

$$P_{X|Y}(x|y) = P[X=x \mid Y=y]$$
$$= \frac{P[X=x, Y=y]}{P[Y=y]}$$
$$= \frac{P_{X,Y}(x,y)}{P_Y(y)}$$

Theorem:
$$P_{X,Y}(x,y) = P_{X|Y}(x|y)P_Y(y) = P_{Y|X}(y|x)P_X(x)$$

Example



Find the conditional PMF of Y given X=x for each $x \in S_X$

Find PMF $P_X(x)$.

$$P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x,y)$$

$$P_X(x) = \begin{cases} 1/4 & x = 1 \\ 1/8+1/8 & x = 2 \\ 1/12+1/12+1/12 & x = 3 \\ 1/16+1/16+1/16+1/16 & x = 4 \\ 0 & \text{Otherwise} \end{cases}$$

$$P_X(x) = \begin{cases} 1/4 & x = 1 \\ 1/4 & x = 2 \\ 1/4 & x = 3 \\ 1/4 & x = 4 \\ 0 & \text{Otherwise} \end{cases}$$

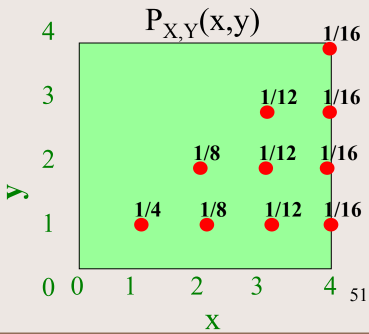
Theorem 4.22 implies that for $x \in \{1,2,3,4\}$

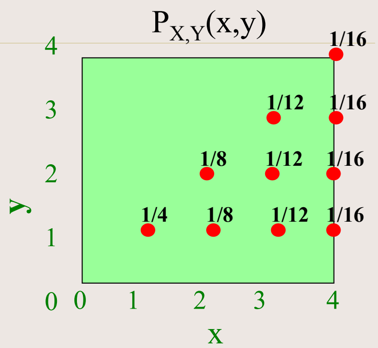
$$P_{Y|X}(y|x) = P_{X,Y}(x,y)/P_X(x) = 4P_{X,Y}(x,y)$$

For each $x \in \{1,2,3,4\}$, $P_{Y|X}(y|x)$ is a different PMF

$$P_{Y|X}(y|1) = \begin{cases} 1 & y = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$P_{Y|X}(y|2) = \begin{cases} \tfrac{1}{2} & y \in \{1,2\} \\ 0 & \text{otherwise} \end{cases}$$





$$P_{Y|X}(y|3) = \begin{cases} 1/3, & y = 1,2,3 \\ 0 & \text{otherwise} \end{cases}$$

$$P_{Y|X}(y|4) = \begin{cases} 1/4, & y = 1,2,3,4 \\ 0 & \text{otherwise} \end{cases}$$

Independent RVs

- From the independent definition
  A and B are independent iff $P[AB] = P[A]P[B]$

- X and Y are independent RVs if and only if
  $\{X=x\}$ and $\{Y=y\}$ are independent for all x,y in $S_{X,Y}$

Definition: $P_{X,Y}(x,y) = P_X(x)P_Y(y)$

Independent RVs

Theorem: If X and Y are statistically independent,

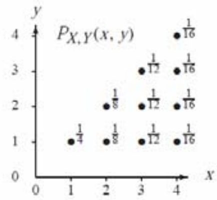
(a) $r_{X,Y} = E[XY] = E[X]E[Y]$
(b) $E[X|Y = y] = E[X]$ for all $y \in S_Y$
(c) $E[Y|X = x] = E[Y]$ for all $x \in S_X$
(d) $Var[X+Y] = Var[X] + Var[Y]$
(e) $Cov[X,Y] = \rho_{X,Y} = 0$

More than 2 RVs

Definition: Joint PMF of discrete RV $X_1,\ldots X_N$ is

$$P_{X_1,\ldots,X_N}(x_1,\ldots,x_N) = P[X_1=x_1, \ldots , X_N=x_N]$$

Example 4.17



Random variables X and Y have the joint PMF $P_{X,Y}(x, y)$, as given in Example 4.13 and repeated in the accompanying graph. Find the conditional PMF of Y given $X = x$ for each $x \in S_X$.

To apply Theorem 4.22, we first find the marginal PMF $P_X(x)$. By Theorem 4.3, $P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x, y)$. For a given $X = x$, we sum the nonzero probablities along the vertical line $X = x$. That is,

$$P_X(x) = \begin{cases} 1/4 & x = 1, \\ 1/8 + 1/8 & x = 2, \\ 1/12 + 1/12 + 1/12 & x = 3, \\ 1/16 + 1/16 + 1/16 + 1/16 & x = 4, \\ 0 & \text{otherwise,} \end{cases} = \begin{cases} 1/4 & x = 1, \\ 1/4 & x = 2, \\ 1/4 & x = 3, \\ 1/4 & x = 4, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 4.22 implies that for $x \in \{1, 2, 3, 4\}$,

$$P_{Y|X}(y|x) = \frac{P_{X,Y}(x, y)}{P_X(x)} = 4P_{X,Y}(x, y). \tag{4.98}$$

For each $x \in \{1, 2, 3, 4\}$, $P_{Y|X}(y|x)$ is a different PMF.

$$P_{Y|X}(y|1) = \begin{cases} 1 & y = 1, \\ 0 & \text{otherwise.} \end{cases} \qquad P_{Y|X}(y|2) = \begin{cases} 1/2 & y \in \{1, 2\}, \\ 0 & \text{otherwise.} \end{cases}$$

$$P_{Y|X}(y|3) = \begin{cases} 1/3 & y \in \{1, 2, 3\}, \\ 0 & \text{otherwise.} \end{cases} \qquad P_{Y|X}(y|4) = \begin{cases} 1/4 & y \in \{1, 2, 3, 4\}, \\ 0 & \text{otherwise.} \end{cases}$$

Given $X = x$, the conditional PMF of Y is the discrete uniform $(1, x)$ random variable.

Example 4.25 Random variables X and Y have a joint PMF given by the following matrix

$P_{X,Y}(x,y)$	$y=-1$	$y=0$	$y=1$
$x=-1$	0	0.25	0
$x=1$	0.25	0.25	0.25

(4.140)

Are X and Y independent? Are X and Y uncorrelated?

For the marginal PMFs, we have $P_X(-1) = 0.25$ and $P_Y(-1) = 0.25$. Thus

$$P_X(-1)\,P_Y(-1) = 0.0625 \neq P_{X,Y}(-1,-1) = 0, \tag{4.141}$$

and we conclude that X and Y are not independent.

To find $\text{Cov}[X,Y]$, we calculate

$$E[X] = 0.5, \qquad E[Y] = 0, \qquad E[XY] = 0. \tag{4.142}$$

Therefore, Theorem 4.16(a) implies

$$\text{Cov}[X,Y] = E[XY] - E[X]\,E[Y] = \rho_{X,Y} = 0, \tag{4.143}$$

and by definition X and Y are uncorrelated.

Example 4.18 In Example 4.17, we derived the following conditional PMFs: $P_{Y|X}(y|1)$, $P_{Y|X}(y|2)$, $P_{Y|X}(y|3)$, and $P_{Y|X}(y|4)$. Find $E[Y|X=x]$ for $x = 1, 2, 3, 4$.

Applying Theorem 4.23 with $g(x,y) = x$, we calculate

$$E[Y|X=1] = 1, \qquad E[Y|X=2] = 1.5, \tag{4.100}$$

$$E[Y|X=3] = 2, \qquad E[Y|X=4] = 2.5. \tag{4.101}$$

Now we consider the case in which X and Y are continuous random variables. We observe $\{Y = y\}$ and define the PDF of X given $\{Y = y\}$. We cannot use $B = \{Y = y\}$ in Definition 4.10 because $P[Y = y] = 0$. Instead, we define a *conditional probability density function*, denoted as $f_{X|Y}(x|y)$.