

Joint Distributions

CSE 312 Su23
Lecture 16

Today

A closer look at more interesting probabilistic models!

Thus far, we've seen some basic models. A series of coin flips, car accidents, and the like.

Most interesting probabilistic models involve more than one random variable.

Modelling multiple RVs

Some common situations:

medical history, symptoms, personal information => predict illness type

traffic, pedestrians, weather, traffic lights => response of self-driving car

But wait!

We have looked at multiple random variables! We've looked at independence and such!

You're right.

But what happens when we're modelling multiple **dependent** random variables?

We'll see this over the next few days...

Joint PMFs and Joint Range

For two (discrete) random variables X, Y , their joint PMF is:

$$p_{X,Y}(a, b) = \mathbb{P}(X = a \cap Y = b) = \mathbb{P}(X = a, Y = b)$$

The joint range of $p_{X,Y}$ is

$$\Omega_{X,Y} = \{(c, d) : p_{X,Y}(c, d) > 0\} \subseteq \Omega_X \times \Omega_Y$$

Note:

$$\sum_{(s,t) \in \Omega_{X,Y}} p_{X,Y}(s, t) = 1$$

Examples

Roll a blue die and a red die. Each die is 4-sided. Let X be the blue die's result and Y be the red die's result.

Each die (individually) is fair. But not all results are equally likely when looking at them both together.

$$p_{X,Y}(1,2) = 3/16.$$

$p_{X,Y}$	$X=1$	$X=2$	$X=3$	$X=4$
$Y=1$	1/16	1/16	1/16	1/16
$Y=2$	3/16	0	0	1/16
$Y=3$	0	2/16	0	2/16
$Y=4$	0	1/16	3/16	0

* Table of all possible outcomes for several discrete RVs.

Examples

$$\Omega_{X,Y} = \{(c, d) : p_{X,Y}(c, d) > 0\} \subseteq \Omega_X \times \Omega_Y$$

$$\Omega_{X,Y} = \{(1,1), (1,2), (2,1), (2,3), (2,4), (3,1), (3,4), (4,1), (4,2), (4,3)\}$$

$$\sum_{(s,t) \in \Omega_{X,Y}} p_{X,Y}(s, t) = 1$$

$$\frac{1 + 1 + 1 + 1 + 3 + 1 + 2 + 2 + 1 + 3}{16} = 1$$

$p_{X,Y}$	$X=1$	$X=2$	$X=3$	$X=4$
$Y=1$	1/16	1/16	1/16	1/16
$Y=2$	3/16	0	0	1/16
$Y=3$	0	2/16	0	2/16
$Y=4$	0	1/16	3/16	0

Examples

Roll a blue die and a red die. Each die is 4-sided. Let X be the blue die's result and Y be the red die's result.

Each die (individually) is fair. But not all results are equally likely when looking at them both together.

$p_{X,Y}$	$X=1$	$X=2$	$X=3$	$X=4$
$Y=1$	1/16	1/16	1/16	1/16
$Y=2$	3/16	0	0	1/16
$Y=3$	0	2/16	0	2/16
$Y=4$	0	1/16	3/16	0

Marginal Distributions

What if I just want to talk about X ?

Well, use the law of total probability:

$$\mathbb{P}(X = k) = \sum_{\text{partition } \{E_i\}} \mathbb{P}(X = k \cap E_i)$$

and let E_i be possible outcomes for Y .

For the dice example

$$\mathbb{P}(X = k) = \sum_{\ell=1}^4 \mathbb{P}(X = k \cap Y = \ell)$$

$$p_X(k) = \sum_{\ell=1}^4 p_{X,Y}(k, \ell)$$

Marginal Distributions

$$p_X(k) = \sum_{\ell=1}^4 p_{X,Y}(k, \ell)$$

So

$$p_X(2) = \frac{1}{16} + 0 + \frac{2}{16} + \frac{1}{16} = \frac{4}{16}$$

$p_{X,Y}$	$X=1$	$X=2$	$X=3$	$X=4$
$Y=1$	1/16	1/16	1/16	1/16
$Y=2$	3/16	0	0	1/16
$Y=3$	0	2/16	0	2/16
$Y=4$	0	1/16	3/16	0

Marginal Distributions

In general, the marginal distributions of the joint PMF are defined as:

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

$$p_Y(b) = \mathbb{P}(Y = b) = \sum_x p_{X,Y}(x, b)$$

$p_X(k)$ is called the “marginal” distribution for X (because we “marginalized” Y) it’s the same PMF we’ve always used; the name emphasizes we have gotten rid of one of the variables.

Weird dice

Roll two fair dice independently.
Let X be the value of the first die,
and Y be the value of the second
die.

Let U be the minimum of the two
rolls (i.e., $U = \min(X, Y)$) and V be
the maximum (i.e., $V = \max(X, Y)$)

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$				
$V=2$				
$V=3$				
$V=4$				

Weird dice

Let U be the minimum of the two rolls (i.e., $U = \min(X, Y)$) and V be the maximum (i.e., $V = \max(X, Y)$)

What is the joint distribution of U and V , $p_{U,V}(u, v)$?

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$				
$V=2$				
$V=3$				
$V=4$				

Weird dice

Let U be the minimum of the two rolls (i.e., $U = \min(X, Y)$) and V be the maximum (i.e., $V = \max(X, Y)$)

What is the joint distribution of U and V , $p_{U,V}(u, v)$?

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$	1/16	0	0	0
$V=2$	2/16	1/16	0	0
$V=3$	2/16	2/16	1/16	0
$V=4$	2/16	2/16	2/16	1/16

Weird dice

Let U be the minimum of the two rolls
(i.e., $U = \min(X, Y)$) and
 V be the maximum (i.e., $V = \max(X, Y)$)

What is the joint distribution of
 U and V , $p_{U,V}(u, v)$?

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$	1/16	0	0	0
$V=2$	2/16	1/16	0	0
$V=3$	2/16	2/16	1/16	0
$V=4$	2/16	2/16	2/16	1/16

$$p_{U,V}(u, v) = \begin{cases} 1/16 & (u, v) \in \Omega_U \times \Omega_V \text{ where } u = v \\ 2/16 & (u, v) \in \Omega_U \times \Omega_V \text{ where } v > u \\ 0 & \text{otherwise} \end{cases}$$

Weird dice

Let U be the minimum of the two rolls
(i.e., $U = \min(X, Y)$) and
 V be the maximum (i.e., $V = \max(X, Y)$)

What's $p_U(z)$? (the marginal for U)

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$	1/16	0	0	0
$V=2$	2/16	1/16	0	0
$V=3$	2/16	2/16	1/16	0
$V=4$	2/16	2/16	2/16	1/16

Weird dice

Let U be the minimum of the two rolls
(i.e., $U = \min(X, Y)$) and
 V be the maximum (i.e., $V = \max(X, Y)$)

What's $p_U(z)$? (the marginal for U)

$$p_U(z) = \begin{cases} \frac{7}{16} & \text{if } z = 1 \\ \frac{5}{16} & \text{if } z = 2 \\ \frac{3}{16} & \text{if } z = 3 \\ \frac{1}{16} & \text{if } z = 4 \\ 0 & \text{otherwise} \end{cases}$$

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$	1/16	0	0	0
$V=2$	2/16	1/16	0	0
$V=3$	2/16	2/16	1/16	0
$V=4$	2/16	2/16	2/16	1/16

Weird dice

Let U be the minimum of the two rolls
(i.e., $U = \min(X, Y)$) and
 V be the maximum (i.e., $V = \max(X, Y)$)

Are U and V independent?

Recall: Discrete RVs U, V are independent iff $p_{U,V}(u, v) = p_U(u) \cdot p_V(v)$
for all $u \in \Omega_U, v \in \Omega_V$

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$	1/16	0	0	0
$V=2$	2/16	1/16	0	0
$V=3$	2/16	2/16	1/16	0
$V=4$	2/16	2/16	2/16	1/16

Weird dice

Are U and V independent?

Recall: Discrete RVs U, V are independent iff $p_{U,V}(u, v) = p_U(u) \cdot p_V(v)$ for all $u \in \Omega_U, v \in \Omega_V$.

I *could* compute the marginal PMF of V , and then iterate through every pair of values of u and v , and check if these equations match.

But I am a bit lazy...

A check for independence

The check: $\Omega_X \times \Omega_Y$ must equal $\Omega_{X,Y}$ for independence.

Suppose that there is some $(a, b) \in \Omega_X \times \Omega_Y$, but not in $\Omega_{X,Y}$.

Then:

$$p_{U,V}(a, b) = 0$$

But:

$$p_U(a) > 0 \text{ and } p_U(b) > 0$$

A check for independence

Suppose that there is some $(a, b) \in \Omega_X \times \Omega_Y$, but not in $\Omega_{X,Y}$.

Then: $p_{U,V}(a, b) = 0$ But: $p_U(a) > 0$ and $p_U(b) > 0$

Discrete RVs U, V are independent iff $p_{U,V}(u, v) = p_U(u) \cdot p_V(v)$ for all $u \in \Omega_U, v \in \Omega_V$.

But the supposed event violates independence.

Therefore, a **necessary but not sufficient condition** for independence:
 $\Omega_X \times \Omega_Y = \Omega_{X,Y}$.

Weird dice

Let U be the minimum of the two rolls
(i.e., $U = \min(X, Y)$) and
 V be the maximum (i.e., $V = \max(X, Y)$)

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$	1/16	0	0	0
$V=2$	2/16	1/16	0	0
$V=3$	2/16	2/16	1/16	0
$V=4$	2/16	2/16	2/16	1/16

Are U and V independent?

A necessary but not sufficient condition for independence: $\Omega_X \times \Omega_Y = \Omega_{X,Y}$.

U and V are not independent.

Joint PDFs...?

Yup.

A joint PDF for two continuous RVs X and Y is a function $f_{X,Y}(x, y)$ defined on $\mathbb{R} \times \mathbb{R}$ such that....

$$f_{X,Y}(x, y) \geq 0 \text{ for all } x, y \in \mathbb{R}$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$$

$$\text{Joint range: } \Omega_{X,Y} = \{(c, d) : f_{X,Y}(c, d) > 0\} \subseteq \Omega_X \times \Omega_Y$$

Joint continuous probabilities

$$\mathbb{P}(a \leq X \leq b \cap c \leq Y \leq d) = \int_a^b \int_c^d f_{X,Y}(x, y) \, dy \, dx$$

Darts

Suppose X and Y are jointly and uniformly distributed on a circle of radius R centered at the origin.

Darts

Suppose X and Y are jointly and uniformly distributed on a circle of radius R centered at the origin.

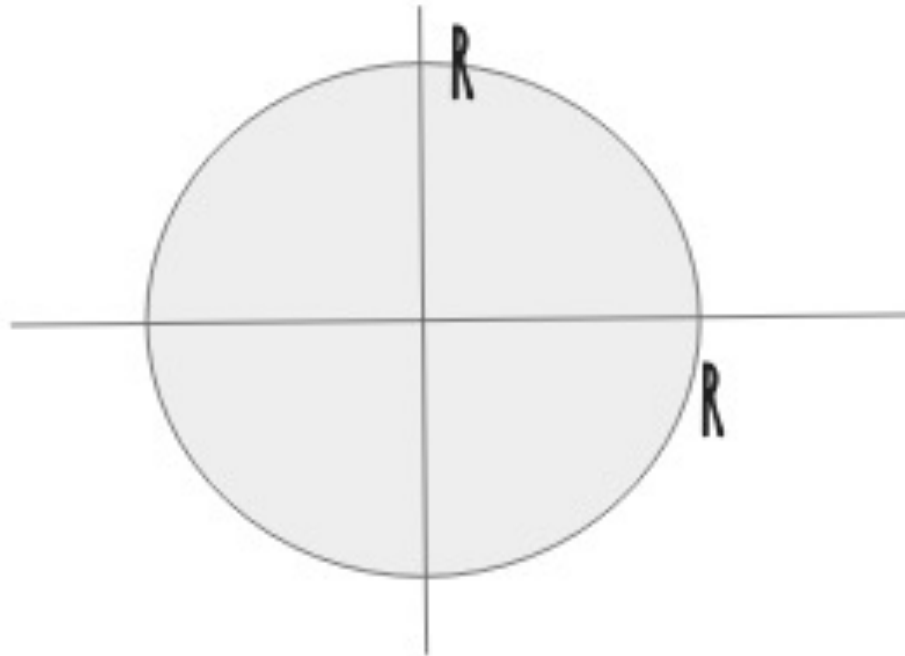
Let's find and sketch the joint range $\Omega_{X,Y}$.

Darts

Suppose X and Y are jointly and uniformly distributed on a circle of radius R centered at the origin.

Let's find and sketch the joint range $\Omega_{X,Y}$.

$$\Omega_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2\}$$



Darts

Suppose X and Y are jointly and uniformly distributed on a circle of radius R centered at the origin.

$$\Omega_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2\}$$

What is the joint PDF?

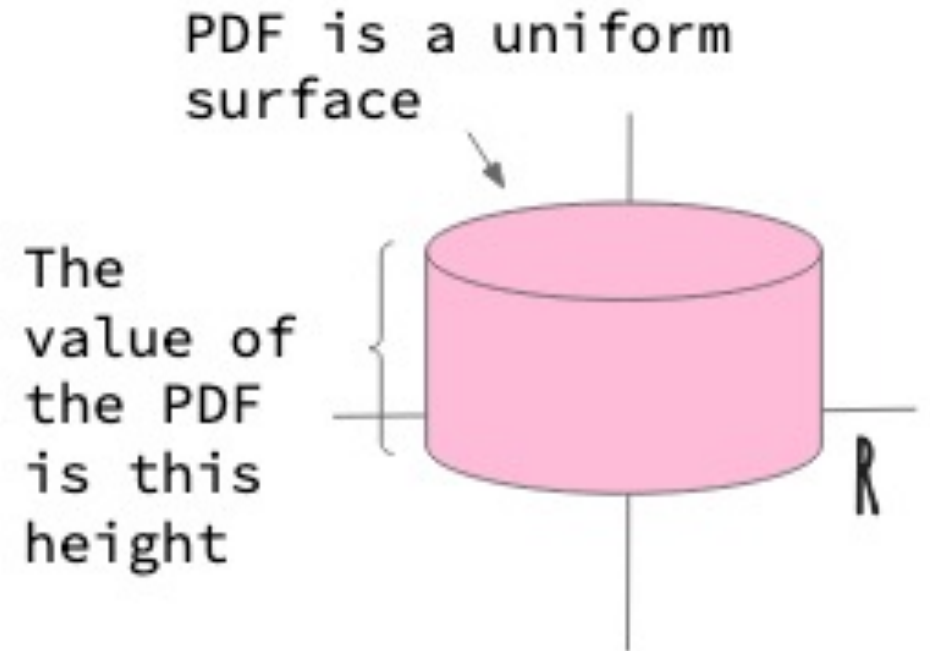
Darts

Suppose X and Y are jointly and uniformly distributed on a circle of radius R centered at the origin.

$$\Omega_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2\}$$

What is the joint PDF?

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{\pi R^2} & x, y \in \Omega_{X,Y} \\ 0 & \text{otherwise} \end{cases}$$



Darts

Suppose X and Y are jointly and uniformly distributed on a circle of radius R centered at the origin.

$$\Omega_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2\}$$

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{\pi R^2} & x, y \in \Omega_{X,Y} \\ 0 & \text{otherwise} \end{cases}$$

What about Ω_X and $f_X(x)$?

Darts

Suppose X and Y are jointly and uniformly distributed on a circle of radius R centered at the origin.

$$\Omega_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2\}$$

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{\pi R^2} & x, y \in \Omega_{X,Y} \\ 0 & \text{otherwise} \end{cases}$$

What about Ω_X and $f_X(x)$?

$$\Omega_X = [-R, R]$$

$$f_X(x) = \int_{-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} f_{X,Y}(x, y) dy$$

Darts

Suppose X and Y are jointly and uniformly distributed on a circle of radius R centered at the origin.

$$\Omega_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2\}$$

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{\pi R^2} & x, y \in \Omega_{X,Y} \\ 0 & \text{otherwise} \end{cases}$$

Are X and Y independent?

Darts

Suppose X and Y are jointly and uniformly distributed on a circle of radius R centered at the origin.

$$\Omega_{X,Y} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2\}$$

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{\pi R^2} & x, y \in \Omega_{X,Y} \\ 0 & \text{otherwise} \end{cases}$$

Are X and Y independent? No, $\Omega_{X,Y} \neq \Omega_X \times \Omega_Y$.

Continuous Servers

The time until server 1 crashes is $X \sim \text{Exp}(\lambda)$ and the time until server 2 crashes (independent of server 1) is $Y \sim \text{Exp}(\mu)$.

What is the probability server 1 crashes before server 2?

$$\mathbb{P}(X < Y)$$

Continuous Servers

The time until server 1 crashes is $X \sim \text{Exp}(\lambda)$ and the time until server 2 crashes (independent of server 1) is $Y \sim \text{Exp}(\mu)$.

What is the probability server 1 crashes before server 2?

$$\mathbb{P}(X < Y)$$

$$= \int_0^\infty \int_x^\infty f_{X,Y}(x, y) dy dx$$

$$= \int_0^\infty \int_x^\infty f_X(x) \cdot f_Y(y) dy dx$$

Law of Total Probability (Continuous)

Let A be an event and Y a continuous random variable, then

$$\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A \mid Y = y) f_Y(y) dy$$

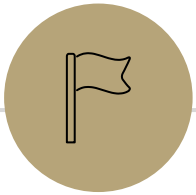
Joint Expectation

Expectations of joint functions

For a function $g(X, Y)$, the expectation can be written in terms of the joint pmf.

$$\mathbb{E}[g(X, Y)] = \sum_{x \in \Omega_X} \sum_{y \in \Omega_Y} g(x, y) \cdot p_{XY}(x, y)$$

This definition hopefully isn't surprising at this point (it's the value of g times the probability g takes on that value), but it's good to have down



Conditional Distributions

A set of more formulae.

But really, explicitly just shifting our knowledge of conditional probabilities of events to random variables... to build up to an elegant result 😊

Conditional PMFs/PDFs

Conditional PMF:

$$p_{X|Y}(a | b) = \mathbb{P}(X = a | Y = b) = \frac{p_{X,Y}(a,b)}{p_Y(b)} = \frac{p_{Y|X}(b | a) p_X(a)}{p_Y(b)}$$

Conditional PDF:

$$f_{X|Y}(a | b) = \frac{f_{X,Y}(a,b)}{f_Y(b)} = \frac{f_{Y|X}(b | a) f_X(a)}{f_Y(b)}$$

Conditional Expectation

Waaaaaay back when, we said conditioning on an event creates a new probability space, with all the laws holding.

So, we can define things like “conditional expectations” which is the expectation of a random variable in that new probability space.

$$\mathbb{E}[X|A] = \sum_{x \in \Omega} x \cdot \mathbb{P}(X = x|A)$$

$$\mathbb{E}[X|Y = y] = \sum_{x \in \Omega_X} x \cdot \mathbb{P}(X = x|Y = y)$$

Conditional Expectations

All your favorite theorems are still true.

For example, linearity of expectation still holds

$$\mathbb{E}[(aX + bY + c) | A] = a\mathbb{E}[X|A] + b\mathbb{E}[Y|A] + c$$

Law of Total Expectation

Let $A_1, A_2, \dots, A_k$ be a partition of the sample space, then

$$\mathbb{E}[X] = \sum_{i=1}^k \mathbb{E}[X|A_i] \mathbb{P}(A_i)$$

Similar in form to law of total probability, and the proof goes that way as well.

Law of Total Expectation

Let $A_1, A_2, \dots, A_k$ be a partition of the sample space, then

$$\mathbb{E}[X] = \sum_{i=1}^k \mathbb{E}[X|A_i] \mathbb{P}(A_i)$$

Let X, Y be discrete random variables, then

$$\mathbb{E}[X] = \sum_{y \in \Omega_Y} \mathbb{E}[X|Y = y] \mathbb{P}(Y = y)$$

Similar in form to law of total probability, and the proof goes that way as well.

LTE

You will flip 2 (independent, fair coins). Call the number of heads X . Then (independently of the coin flips) draw a geometric random variable Y from the distribution $\text{Geo}(X + 1)$.

What is $\mathbb{E}[Y]$?

LTE

You will flip 2 (independent, fair coins). Call the number of heads X . Then (independently of the coin flips) draw a geometric random variable Y from the distribution $Geo(X + 1)$.

What is $\mathbb{E}[Y]$?

$$\mathbb{E}[Y]$$

$$= \mathbb{E}[Y|X = 0]\mathbb{P}(X = 0) + \mathbb{E}[Y|X = 1]\mathbb{P}(X = 1) + \mathbb{E}[Y|X = 2]\mathbb{P}(X = 2)$$

$$= \mathbb{E}[Y|X = 0] \cdot \frac{1}{4} + \mathbb{E}[Y|X = 1] \cdot \frac{1}{2} + \mathbb{E}[Y|X = 2] \cdot \frac{1}{4}$$

$$= \frac{1}{0+1} \cdot \frac{1}{4} + \frac{1}{1+1} \cdot \frac{1}{2} + \frac{1}{2+1} \cdot \frac{1}{4} = \frac{7}{12}.$$

Elevator Rides

X people enter an elevator on the ground floor, where $X \sim Poi(10)$.

There are N floors above the ground floor, and each person is equally likely to get off at any one of the N floors (independent of where others get off).

What is the expected number of stops the elevator makes before all passengers get off?

Elevator Rides

X people enter an elevator on the ground floor, where $X \sim \text{Poi}(10)$. There are N floors above the ground floor, and each person is equally likely to get off at any one of the N floors (independent of where others get off). What is the expected number of stops the elevator makes before all passengers get off?

Elevator Rides

X people enter an elevator on the ground floor, where $X \sim \text{Poi}(10)$. There are N floors above the ground floor, and each person is equally likely to get off at any one of the N floors (independent of where others get off). What is the expected number of stops the elevator makes before all passengers get off?

Analogues for continuous

Everything we saw today has a continuous version.

There are “no surprises”– replace pmf with pdf and sums with integrals.

	Discrete	Continuous
Joint PMF/PDF	$p_{X,Y}(x, y) = P(X = x, Y = y)$	$f_{X,Y}(x, y) \neq P(X = x, Y = y)$
Joint CDF	$F_{X,Y}(x, y) = \sum_{t \leq x} \sum_{s \leq y} p_{X,Y}(t, s)$	$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(t, s) ds dt$
Normalization	$\sum_x \sum_y p_{X,Y}(x, y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x, y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$
Expectation	$E[g(X, Y)] = \sum_x \sum_y g(x, y) p_{X,Y}(x, y)$	$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy$
Conditional PMF/PDF	$p_{X Y}(x y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$	$f_{X Y}(x y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$
Conditional Expectation	$E[X Y = y] = \sum_x x p_{X Y}(x y)$	$E[X Y = y] = \int_{-\infty}^{\infty} x f_{X Y}(x y) dx$
Independence	$\forall x, y, p_{X,Y}(x, y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x, y) = f_X(x) f_Y(y)$