



# Mathematical Foundations

Data Science: Jordan Boyd-Graber  
University of Maryland

SLIDES ADAPTED FROM DAVE BLEI AND LAUREN HANNAH

## Marginalization

If we know a joint distribution of multiple variables, what if we want to know the distribution of only one of the variables?

We can compute the distribution of  $P(X)$  from  $P(X, Y, Z)$  through *marginalization*:

$$\begin{aligned}\sum_y \sum_z P(X, Y = y, Z = z) &= \sum_y \sum_z P(X)P(Y = y, Z = z | X) \\ &= P(X) \sum_y \sum_z P(Y = y, Z = z | X) \\ &= P(X)\end{aligned}$$

## Marginalization

If we know a joint distribution of multiple variables, what if we want to know the distribution of only one of the variables?

We can compute the distribution of  $P(X)$  from  $P(X, Y, Z)$  through *marginalization*:

$$\begin{aligned}\sum_y \sum_z P(X, Y = y, Z = z) &= \sum_y \sum_z P(X) P(Y = y, Z = z | X) \\ &= P(X) \sum_y \sum_z P(Y = y, Z = z | X) \\ &= P(X)\end{aligned}$$

We'll explain this notation more next week for now the formula is the most important part.

## Marginalization (from Leyton-Brown)

### Joint distribution

temperature (T) and weather (W)

	T=Hot	T=Mild	T=Cold
W=Sunny	.10	.20	.10
W=Cloudy	.05	.35	.20

Marginalization allows us to compute distributions over smaller sets of variables:

- $P(X, Y) = \sum_z P(X, Y, Z = z)$
- Corresponds to summing out a table dimension
- New table still sums to 1

- Marginalize out weather
- Marginalize out temperature

## Marginalization (from Leyton-Brown)

### Joint distribution

temperature (T) and weather (W)

	T=Hot	T=Mild	T=Cold
W=Sunny	.10	.20	.10
W=Cloudy	.05	.35	.20

Marginalization allows us to compute distributions over smaller sets of variables:

- $P(X, Y) = \sum_z P(X, Y, Z = z)$
- Corresponds to summing out a table dimension
- New table still sums to 1

- Marginalize out weather  
T=Hot   T=Mild   T=Cold
- Marginalize out temperature

## Marginalization (from Leyton-Brown)

### Joint distribution

temperature (T) and weather (W)

	T=Hot	T=Mild	T=Cold
W=Sunny	.10	.20	.10
W=Cloudy	.05	.35	.20

Marginalization allows us to compute distributions over smaller sets of variables:

- $P(X, Y) = \sum_z P(X, Y, Z = z)$
- Corresponds to summing out a table dimension
- New table still sums to 1

- Marginalize out weather  
T=Hot   T=Mild   T=Cold
- Marginalize out temperature

## Marginalization (from Leyton-Brown)

### Joint distribution

temperature (T) and weather (W)

	T=Hot	T=Mild	T=Cold
W=Sunny	.10	.20	.10
W=Cloudy	.05	.35	.20

Marginalization allows us to compute distributions over smaller sets of variables:

- $P(X, Y) = \sum_z P(X, Y, Z = z)$
- Corresponds to summing out a table dimension
- New table still sums to 1

- Marginalize out weather  

T=Hot	T=Mild	T=Cold
<hr/>		
.15		
<hr/>		
- Marginalize out temperature

## Marginalization (from Leyton-Brown)

### Joint distribution

temperature (T) and weather (W)

	T=Hot	T=Mild	T=Cold
W=Sunny	.10	.20	.10
W=Cloudy	.05	.35	.20

Marginalization allows us to compute distributions over smaller sets of variables:

- $P(X, Y) = \sum_z P(X, Y, Z = z)$
- Corresponds to summing out a table dimension
- New table still sums to 1

- Marginalize out weather

T=Hot	T=Mild	T=Cold
.15	.55	.30

- Marginalize out temperature

## Marginalization (from Leyton-Brown)

### Joint distribution

temperature (T) and weather (W)

	T=Hot	T=Mild	T=Cold
W=Sunny	.10	.20	.10
W=Cloudy	.05	.35	.20

Marginalization allows us to compute distributions over smaller sets of variables:

- $P(X, Y) = \sum_z P(X, Y, Z = z)$
- Corresponds to summing out a table dimension
- New table still sums to 1

- Marginalize out weather

T=Hot	T=Mild	T=Cold
.15	.55	.30

- Marginalize out temperature

W=Sunny
W=Cloudy

## Marginalization (from Leyton-Brown)

### Joint distribution

temperature (T) and weather (W)

	T=Hot	T=Mild	T=Cold
W=Sunny	.10	.20	.10
W=Cloudy	.05	.35	.20

Marginalization allows us to compute distributions over smaller sets of variables:

- $P(X, Y) = \sum_z P(X, Y, Z = z)$
- Corresponds to summing out a table dimension
- New table still sums to 1

- Marginalize out weather  

T=Hot	T=Mild	T=Cold
.15	.55	.30
- Marginalize out temperature  

W=Sunny
W=Cloudy

## Marginalization (from Leyton-Brown)

### Joint distribution

temperature (T) and weather (W)

	T=Hot	T=Mild	T=Cold
W=Sunny	.10	.20	.10
W=Cloudy	.05	.35	.20

Marginalization allows us to compute distributions over smaller sets of variables:

- $P(X, Y) = \sum_z P(X, Y, Z = z)$
- Corresponds to summing out a table dimension
- New table still sums to 1

- Marginalize out weather  

T=Hot	T=Mild	T=Cold
.15	.55	.30
- Marginalize out temperature  

W=Sunny	.40
W=Cloudy	

## Marginalization (from Leyton-Brown)

### Joint distribution

temperature (T) and weather (W)

	T=Hot	T=Mild	T=Cold
W=Sunny	.10	.20	.10
W=Cloudy	.05	.35	.20

Marginalization allows us to compute distributions over smaller sets of variables:

- $P(X, Y) = \sum_z P(X, Y, Z = z)$
- Corresponds to summing out a table dimension
- New table still sums to 1

- Marginalize out weather

T=Hot	T=Mild	T=Cold
.15	.55	.30

- Marginalize out temperature

W=Sunny	.40
W=Cloudy	.60