

| Coin | Card | Candy | P(Coin, Card, Candy) | |
|-------|--------|-----------|----------------------------|--------------------|
| tails | black | 1 | 0.15 | |
| tails | black | 2 | 0.06 | The probabilities |
| tails | black | 3 | 0.09 | in the last column |
| tails | red | 1 | 0.02 | sum to 1 |
| tails | red | 2 | 0.06 | |
| tails | red | 3 | 0.12 | |
| heads | black | 1 | 0.075 | |
| heads | black | 2 | 0.03 | |
| heads | black | 3 | 0.045 | |
| heads | red | 1 | 0.035 | |
| heads | red | 2 | 0.105 | |
| heads | red | 3 | 0.21 | |
| This | cell m | eans P(Cc | oin=heads, Card=red, Candy | y=3) = 0.21 2 |
| | | | ; | |

Joint Probability Distribution

From the full joint probability distribution, we can calculate any probability involving these three random variables.

e.g. P(*Coin* = *heads* OR *Card* = *red*)

Joint Probability Distribution

P(Coin = heads OR Card = red) =

P(Coin=heads, Card=black, Candy=1) + P(Coin=heads, Card=black, Candy=2) + P(Coin=heads, Card=black, Candy=3) + P(Coin=tails, Card=red, Candy=1) + P(Coin=tails, Card=red, Candy=2) + P(Coin=heads, Card=red, Candy=3) + P(Coin=heads, Card=red, Candy=2) + P(Coin=heads, Card=red, Candy=2) + P(Coin=heads, Card=red, Candy=3) = 0.075 + 0.03 + 0.045 + 0.02 + 0.06 + 0.12 + 0.035 + 0.105 + 0.21 = 0.7

Marginalization

We can even calculate marginal probabilities (the probability distribution over a subset of the variables) e.g.:

P(Coin=tails, Card=red) = P(Coin=tails, Card=red, Candy=1) + P(Coin=tails, Card=red, Candy=2) + P(Coin=tails, Card=red, Candy=3) = 0.02 + 0.06 + 0.12 = 0.2

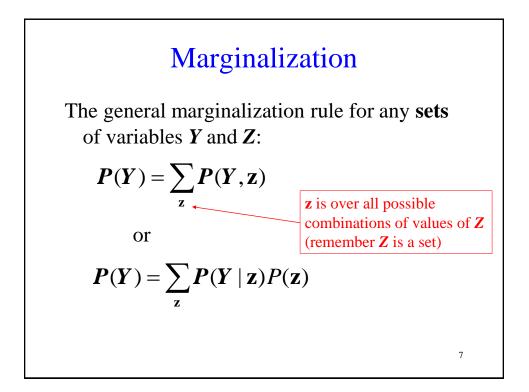
Marginalization

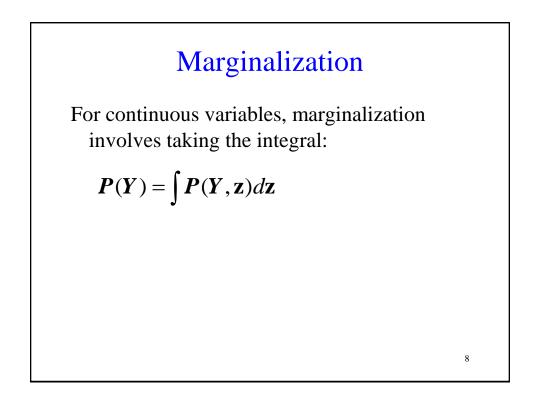
5

6

Or even:

P(Card=black) = P(Coin=heads, Card=black, Candy=1) + P(Coin=heads, Card=black, Candy=2) + P(Coin=heads, Card=black, Candy=3) + P(Coin=tails, Card=black, Candy=1) + P(Coin=tails, Card=black, Candy=2) + P(Coin=tails, Card=black, Candy=3) = 0.075 + 0.03 + 0.045 + 0.015 + 0.06 + 0.09 = 0.315





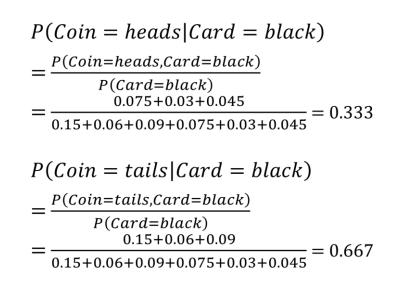
| С | W: | Pra | actic | e |
|-------------------------|-------|-------|-------|----------------------|
| | Coin | Card | Candy | P(Coin, Card, Candy) |
| Compute | tails | black | 1 | 0.15 |
| Compute $P(Candy = 2).$ | tails | black | 2 | 0.06 |
| P(Canay = 2). | tails | black | 3 | 0.09 |
| | tails | red | 1 | 0.02 |
| | tails | red | 2 | 0.06 |
| | tails | red | 3 | 0.12 |
| | heads | black | 1 | 0.075 |
| | heads | black | 2 | 0.03 |
| | heads | black | 3 | 0.045 |
| | heads | red | 1 | 0.035 |
| | heads | red | 2 | 0.105 |
| | heads | red | 3 | 0.21 |
| | | | | 9 |

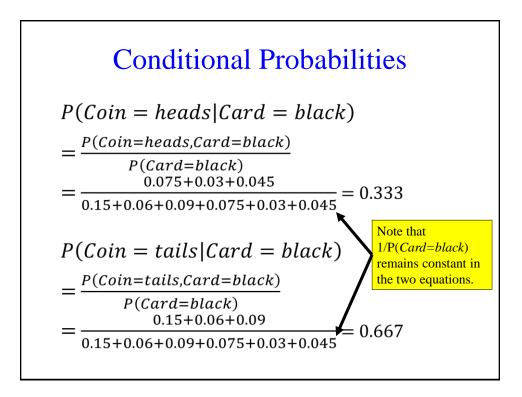
Conditional Probabilities

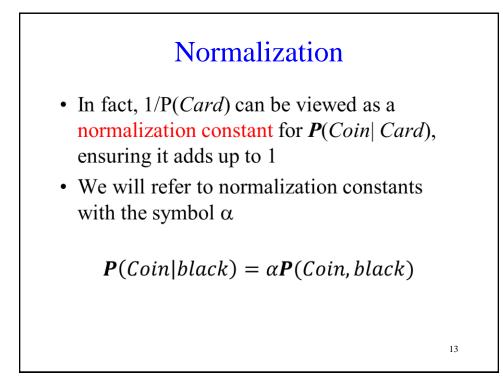
We can also compute conditional probabilities from the joint. Recall:

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

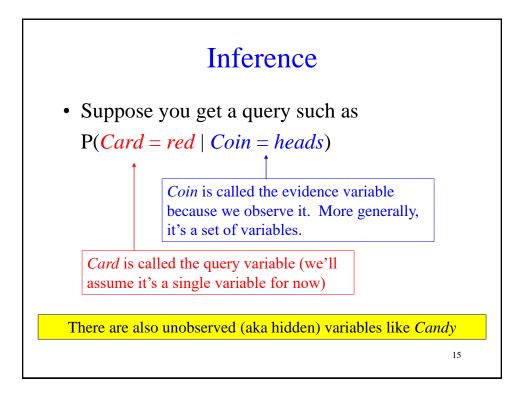
Conditional Probabilities

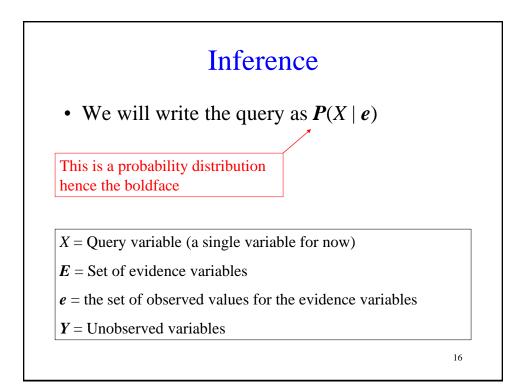


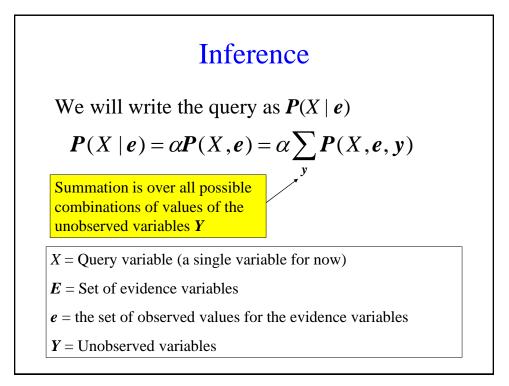


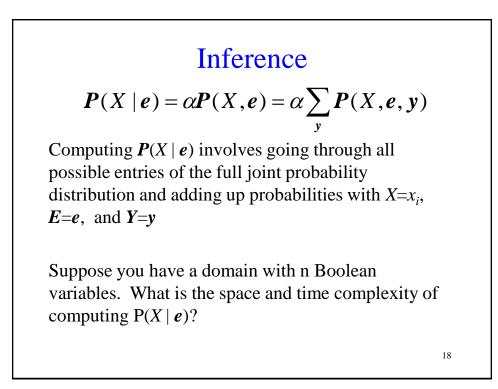


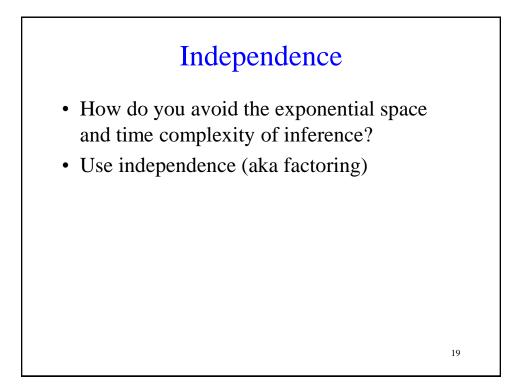
| CW: | Pra | ctio | ce | |
|--------------------------|-------|-------|-------|----------------------|
| | Coin | Card | Candy | P(Coin, Card, Candy) |
| Compute | tails | black | 1 | 0.15 |
| P(Candy = 1 Card = red). | tails | black | 2 | 0.06 |
| | tails | black | 3 | 0.09 |
| | tails | red | 1 | 0.02 |
| | tails | red | 2 | 0.06 |
| | tails | red | 3 | 0.12 |
| | heads | black | 1 | 0.075 |
| | heads | black | 2 | 0.03 |
| | heads | black | 3 | 0.045 |
| | heads | red | 1 | 0.035 |
| | heads | red | 2 | 0.105 |
| | heads | red | 3 | 0.21 |

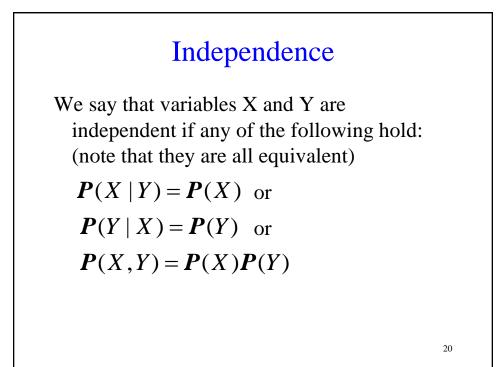


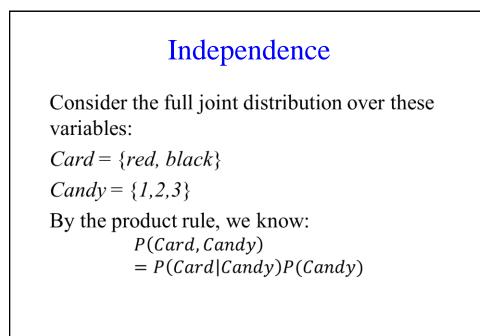


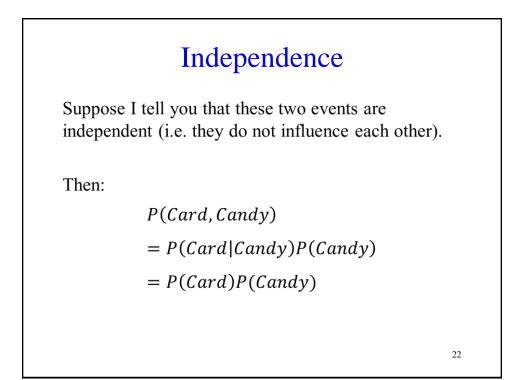


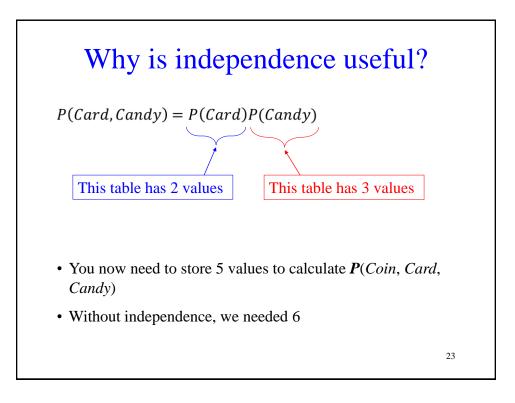


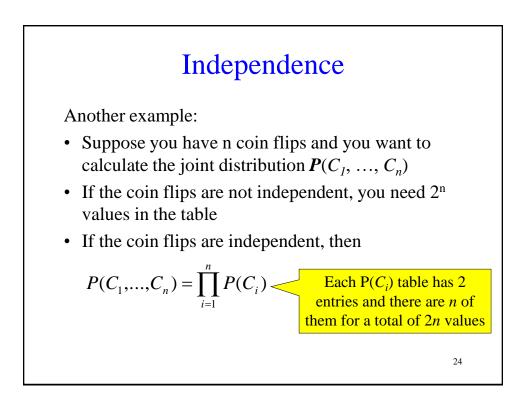












Independence

- Independence is powerful!
- It required extra domain knowledge. A different kind of knowledge than numerical probabilities. It needed an understanding of relationships among the random variables.

25

| | Coin | Card | Candy | P(Coin, Card, Candy) |
|---|---------|-------|-------|----------------------|
| Are Coin and Card | tails | black | 1 | 0.15 |
| • • • • • • • • | tails | black | 2 | 0.06 |
| independent in this | tails | black | 3 | 0.09 |
| distribution? | tails | red | 1 | 0.02 |
| | tails | red | 2 | 0.06 |
| | tails | red | 3 | 0.12 |
| Recall: | heads | black | 1 | 0.075 |
| $\boldsymbol{P}(X \mid Y) = \boldsymbol{P}(X)$ | heads | black | 2 | 0.03 |
| | heads | black | 3 | 0.045 |
| $\boldsymbol{P}(\boldsymbol{Y} \mid \boldsymbol{X}) = \boldsymbol{P}(\boldsymbol{Y})$ | heads | red | 1 | 0.035 |
| $\mathbf{D}(\mathbf{V}, \mathbf{V}) = \mathbf{D}(\mathbf{V}) \mathbf{D}(\mathbf{V})$ | n heads | red | 2 | 0.105 |
| $\boldsymbol{P}(X,Y) = \boldsymbol{P}(X)\boldsymbol{P}(Y)$ | heads | red | 3 | 0.21 |