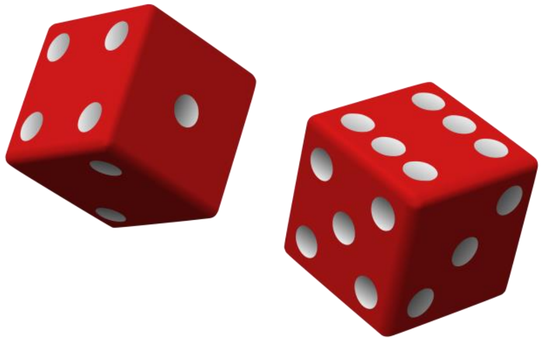
Overview of Probability

Mark Schmidt

September 12, 2017

Practical Application…

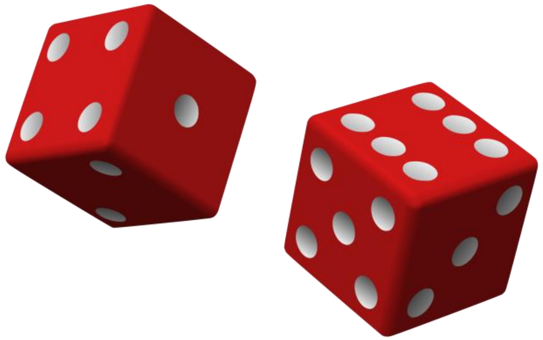
- Dungeons & Dragons scenario:
  - You roll dice 1:
    - Roll 5 or 6 you sneak past monster.
    - Otherwise, you are eaten.
  - If you survive, you roll dice 2:
    - Roll 4-6, find pizza.
    - Otherwise, you find nothing.



https://en.wikipedia.org/wiki/Dice_throw_%28review%29
http://www.dungeonsdragonscartoon.com/2011/11/cloak.html

Practical Application …

- Dungeons & Dragons scenario:
  - You roll dice 1:
    - Roll 5 or 6 you sneak past monster.
    - Otherwise, you are eaten.
  - If you survive, you roll dice 2:
    - Roll 4-6, find pizza.
    - Otherwise, you find nothing.



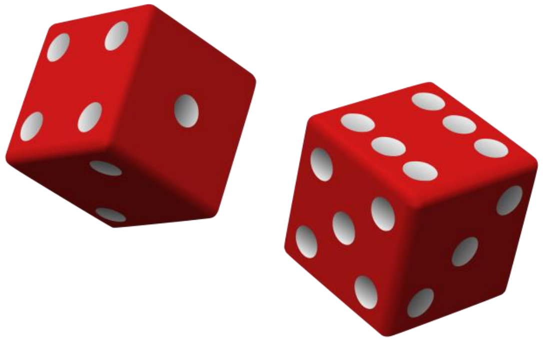
- Probabilities defined on ‘event space’:

D1\D2	1	2	3	4	5	6
1						
2						
3		D_1=3,D_2=2				
4						
5						
6						

https://en.wikipedia.org/wiki/Dice_throw_%28review%29
http://www.dungeonsdragonscartoon.com/2011/11/cloak.html

Practical Application …

- Dungeons & Dragons scenario:
  - You roll dice 1:
    - Roll 5 or 6 you sneak past monster.
    - Otherwise, you are eaten.
  - If you survive, you roll dice 2:
    - Roll 4-6, find pizza.
    - Otherwise, you find nothing.



- Probabilities defined on ‘event space’:

D1\D2	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

¬Survive

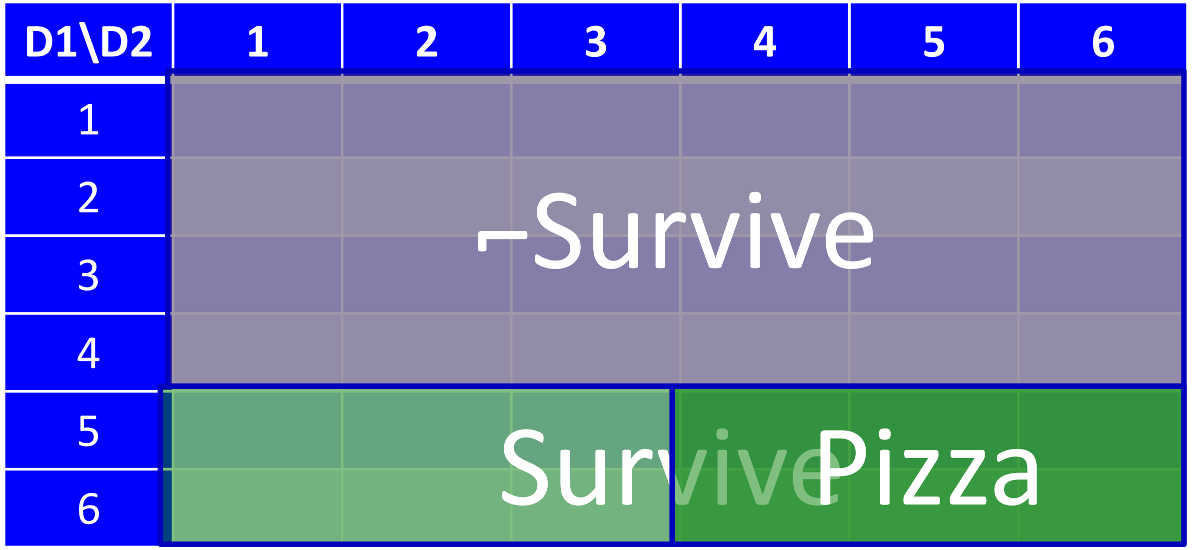
Survive

Pizza

https://en.wikipedia.org/wiki/Dice_throw_%28review%29
http://www.dungeonsdragonscartoon.com/2011/11/cloak.html

Calculating Basic Probabilities

- Probability of event ‘A’ is ratio:
  - p(A) = Area(A)/TotalArea.
  - “Likelihood” that ‘A’ happens.
- Examples:
  - p(Survive) = 12/36 = 1/3.
  - p(Pizza) = 6/36 = 1/6.
  - p(¬Survive) = 1 – p(Survive) = 2/3.



Calculating Basic Probabilities

- Probability of event 'A' is ratio:
  - p(A) = Area(A)/TotalArea.
  - "Likelihood" that 'A' happens.
- Examples:
  - p(Survive) = 12/36 = 1/3.
  - p(Pizza) = 6/36 = 1/6.
  - p(¬Survive) = 1 – p(Survive) = 2/3.
  - p(D_1 is even) = 18/36 = ½.

D1\D2	1	2	3	4	5	6
1						
2	D_1 is even					
3						
4	D_1 is even					
5						
6	D_1 is even					

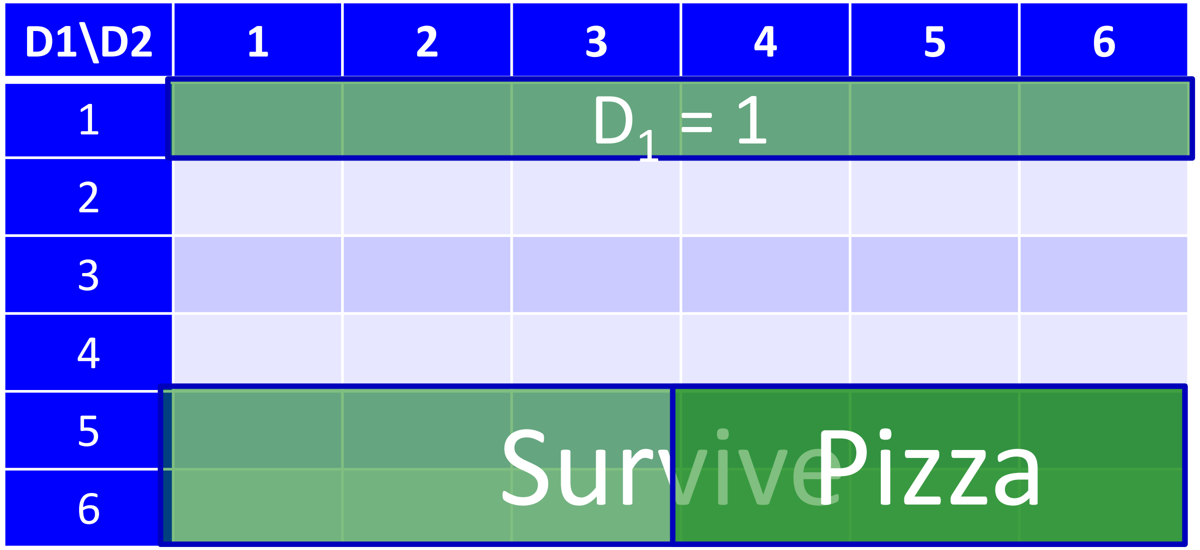
Random Variables and ‘Sum to 1’ Property

- Random variable: variable whose value depends on probability.
- Example: event (D_1 = x) depends on random variable D_1.
- Convention:
  – We’ll use p(x) to mean p(X = x), when random variable X is obvious.
- Sum of probabilities of random variable over entire domain is 1:
  – $\sum_x p(x) = 1$.
  – E.g, $\sum_i p(D_1 = i)$ = 1/6+1/6 + …
  = 1.

<table>
<tr><th>D1\D2</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th></tr>
<tr><td>1</td><td colspan="6">D_1 =1</td></tr>
<tr><td>2</td><td colspan="6">D_1 =2</td></tr>
<tr><td>3</td><td colspan="6">D_1 = 3</td></tr>
<tr><td>4</td><td colspan="6">D_1 = 4</td></tr>
<tr><td>5</td><td colspan="6">D_1 = 5</td></tr>
<tr><td>6</td><td colspan="6">D_1 = 6</td></tr>
</table>

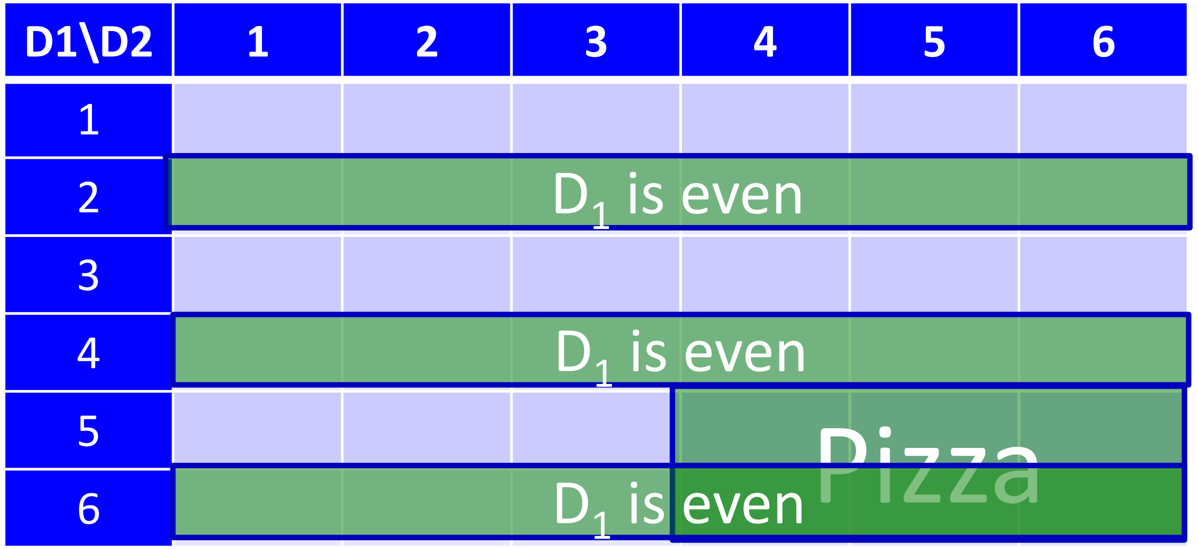
Joint Probability

- Joint probability: probability that A and B happen, written ‘p(A,B)’.
  - Intersection of Area(A) and Area(B).
- Examples:
  - $p(D_1 = 1, \text{Survive}) = 0$.
  - $p(\text{Survive}, \text{Pizza}) = 6/36 = 1/6$.



Joint Probability

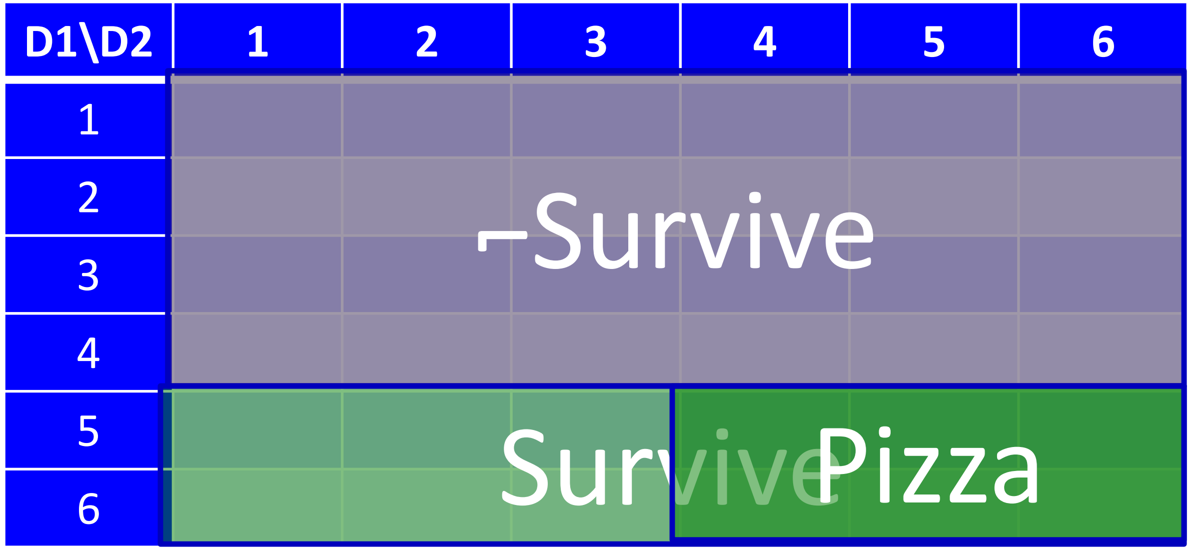
- Joint probability: probability that A and B happen, written ‘p(A,B)’.
  - Intersection of Area(A) and Area(B).
- Examples:
  - $p(D_1 = 1, \text{Survive}) = 0$.
  - $p(\text{Survive}, \text{Pizza}) = 6/36 = 1/6$.
  - $p(D_1 \text{ even}, \text{Pizza}) = 3/36 = 1/12$.



- Note: order of A and B does not matter

Marginalization Rule

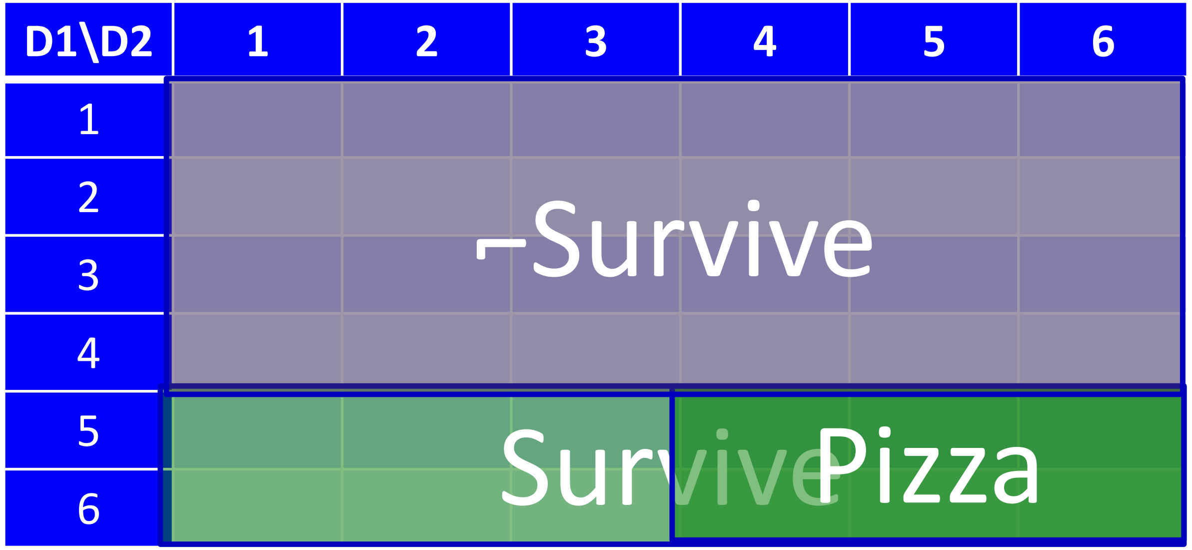
- Marginalization rule:
  - $P(A) = \sum_x P(A, X = x)$.
  - Summing joint over all values of one variable gives probability of the other.
  - Example: $P(Pizza) = P(Pizza, Survive) + P(Pizza, \neg Survive) = \frac{1}{6}$.



  - Applying rule twice: $\sum_x \sum_y p(Y = y, X = x) = 1$.

Conditional Probability

- Conditional probability:
  - probability that A will happen *if we know* that B happens.
  - “probability of A *restricted* to scenarios where B happens”.
  - Written p(A|B), said “probability of A given B”.
- Calculation:
  - Within area of B:
    - Compute Area(A)/TotalArea.
  - p(Pizza | Survive) =

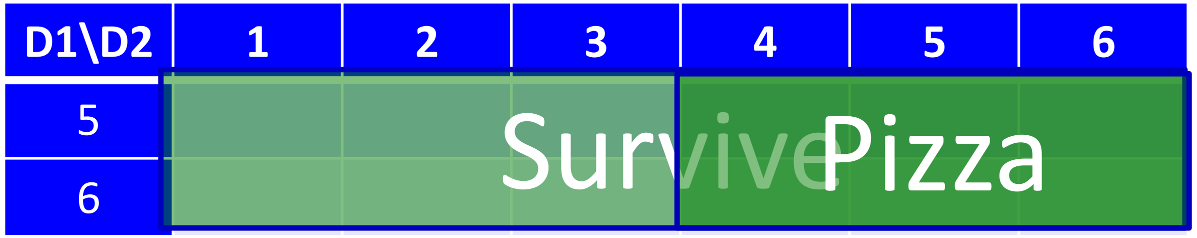


Conditional Probability

- Conditional probability:
  - probability that A will happen *if we know* that B happens.
  - "probability of A *restricted* to scenarios where B happens".
  - Written p(A|B), said "probability of A given B".
- Calculation:
  - Within area of B:
    - Compute Area(A)/TotalArea.
  - p(Pizza | Survive) =

    p(Pizza, Survive)/p(Survive) = 6/12 = ½.
  - Higher than p(Pizza, Survive) = 6/36 = 1/6.
  - More generally, $p(A \mid B) = p(A,B)/p(B)$.

Geometrically: compute area of A on new space where B happened.



‘Sum to 1’ Properties and Bayes Rule.

- Conditional probability P(A | B) sums to one over all A:
  - $\sum_x P(x \mid B) = 1.$
  - P(Pizza | Survive) + P(⌐ Pizza | Survive) = 1.
  - P(Pizza | Survive) + P(Pizza | ⌐Survive) ≠ 1.
- Product rule: p(A,B) = p(A | B)p(B).
- Bayes Rule: $p(A \mid B) = \frac{p(B \mid A)\,p(A)}{p(B)}$
  - Allows you to “reverse” the conditional probability.
- Example:
  - P(Pizza | Survive) = P(Survive | Pizza)P(Pizza)/P(Survive)
    = (1) * (1/6) / (1/3) = ½.
  - http://setosa.io/ev/conditional-probability

Independence of Random Variables

- Events A and B are independent if $p(A,B) = p(A)p(B)$.
  - Equivalently: $p(A|B) = p(A)$.
  - "Knowing B happened tells you nothing about A".
  - We use the notation:

$$A \perp B$$

- Random variables are independent if $p(x,y) = p(x)p(y)$ for all x and y.
  - Flipping two coins:

$$p(C_1 = \text{'heads'}, C_2 = \text{'heads'}) = p(C_1 = \text{'heads'})p(C_2 = \text{'heads'}).$$
$$p(C_1 = \text{'tails'}, C_2 = \text{'heads'}) = p(C_1 = \text{'tails'})p(C_2 = \text{'heads'}).$$

…

Conditional Independence

- A and B are conditionally independent given C if
$$p(A, B \mid C) = p(A \mid C)p(B \mid C).$$
  - Equivalently: $p(A \mid B, C) = p(A \mid C)$.
  - "Knowing C happened, also knowing B happened says nothing about A".
  - Example: $p(\text{Pizza} \mid D_1, \text{Survive}) = p(\text{Pizza} \mid \text{Survive})$.
  - Knowing you survived, dice 1 gives no information about chance of pizza.
  - We use the notation: $A \perp B \mid C$

- Semantics of $p(A, B \mid C, D)$:
  - "probability of A and B happening, if we know that C and D happened".

Conditional Independence

- Example: food poisoning
  - If food was bad, each person independently gets sick with probability 50%
  - Unconditionally, me getting and and you getting sick are NOT independent
    - If I got sick, that makes me think the food was bad, which makes it more likely that you will get sick also. So knowing my situation influences my beliefs about yours.
  - But, conditioned on knowing the food was bad (or not bad), my sickness and your sickness are independent.

More Tutorial Material

- Wikipedia’s conditional probability article is good:
  - https://en.wikipedia.org/wiki/Conditional_probability
- Visual/interactive introduction to probability:
  - http://students.brown.edu/seeing-theory/basic-probability/index.html#first
  - http://students.brown.edu/seeing-theory/compound-probability/index.html#first
- “Probability Primer” (advanced, PP 1.S-5.4 are most relevant):
  - https://www.youtube.com/playlist?list=PL17567A1A3F5DB5E4

Fun with Probabilities

- Probabilities can be used for a huge variety of problems:
  - Are you the hottest person in your group?
  - Poker Odds
  - Are shy students likely to be math students?
  - Battleship
  - Should you put all your eggs/tickets in one basket/lottery?
  - The Price is Right