February 5, 2018

CS 361: Probability & Statistics

Probability

Counting

Multiplication principle

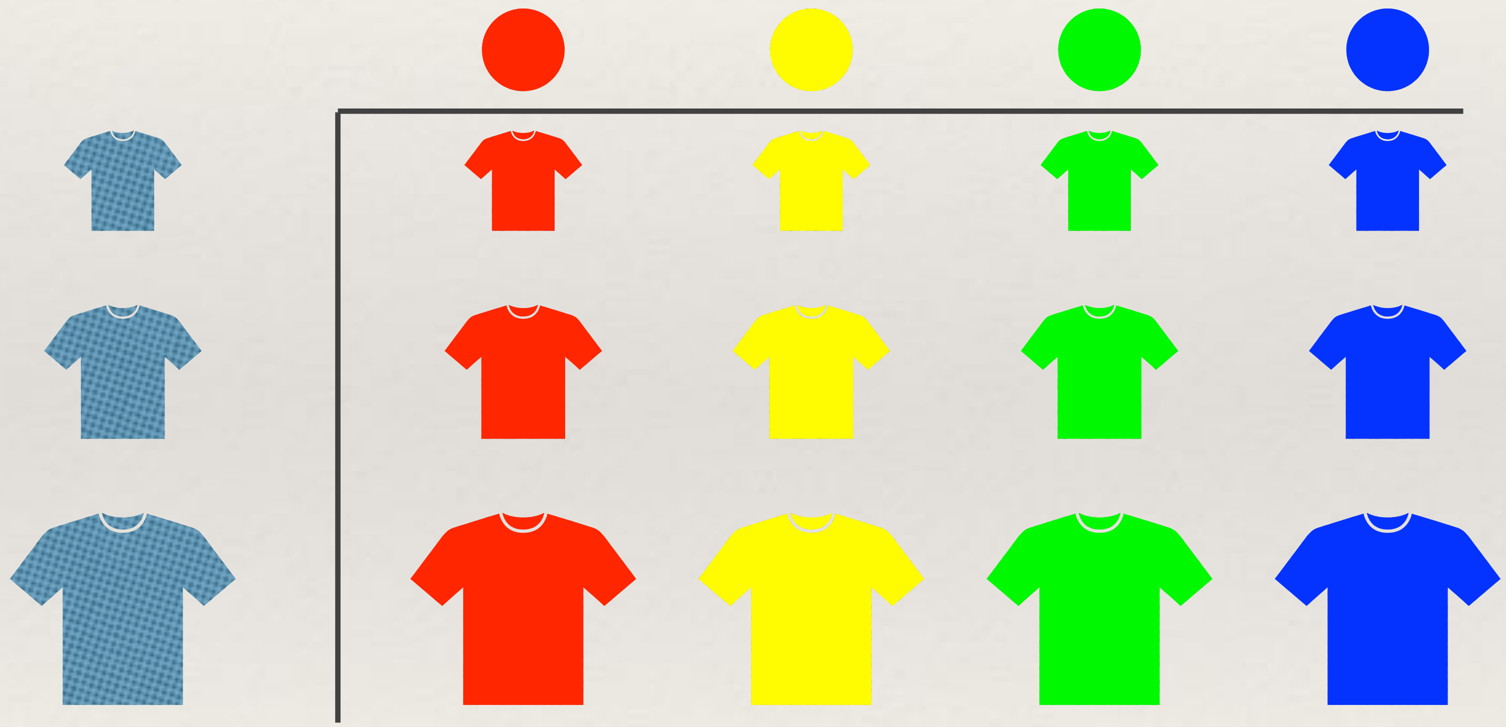
It can be useful to break an experiment down into multiple parts. The number of outcomes is then the product of the number of outcomes of each part

Example: Suppose we grab 3 people at random and ask them what their birthday is. How many possible outcomes might we get?

365 for the first, second, and third person, so 365^3 total possible outcomes

Example

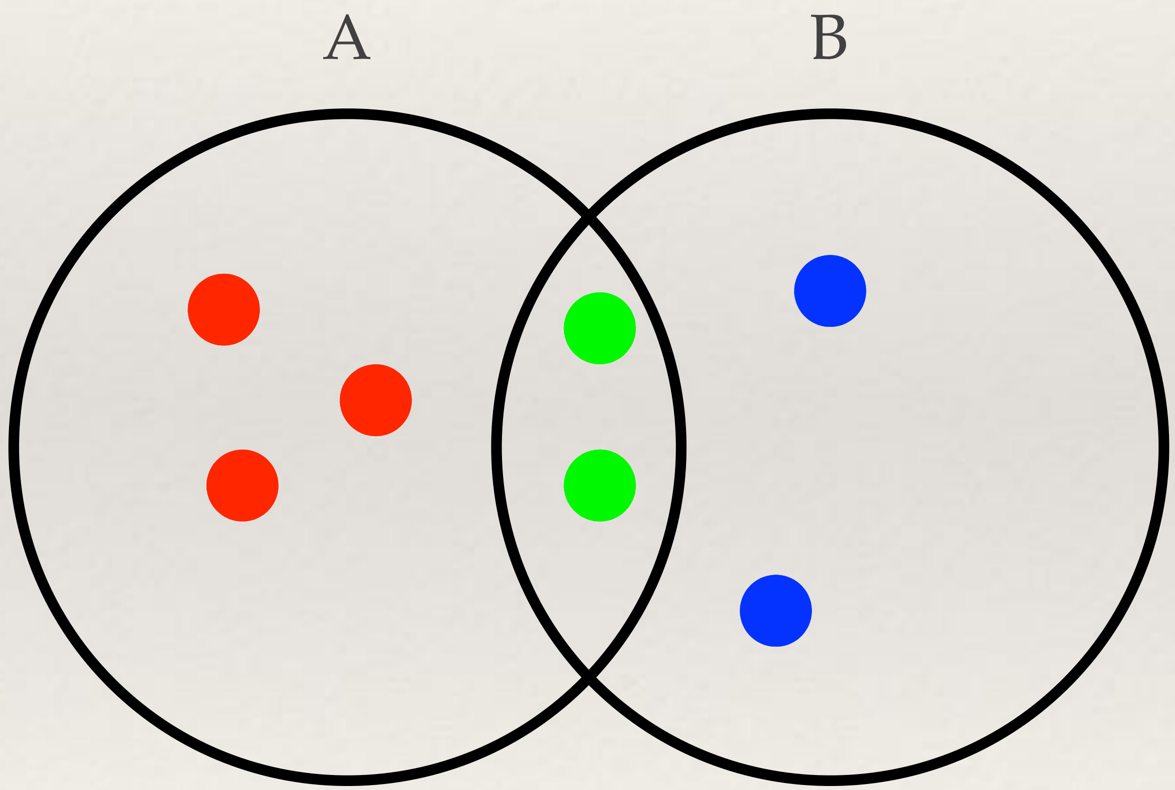
Shirts at a store come in 3 sizes and 4 different colors, how many different shirt types are there?



Inclusion-exclusion principle

Subtract things you would have counted twice

$$|A \cup B| = |A| + |B| - |A \cap B|$$



$$|A \cup B| = |A| + |B| - |A \cap B|$$

5 4 2

Example

License plates in a state can have 5 alpha-numeric characters, how many plates are there that start or end with a number?

A: plates that start with a number

10 x 36 x 36 x 36 x 36 = 16,796,160

B: plates that end with a number

36 x 36 x 36 x 36 x 10 = 16,796,160

We want to know how large A union B is

$$|A \cup B| = |A| + |B| - |A \cap B|$$

A intersect B, starts AND ends with a number

10 x 36 x 36 x 36 x 10 = 4,665,600

A union B contains 16,796,160 + 16,796,160 - 4,665,600 = 28,926,720 plates

Permutations

How many possible re-arrangements are there for the string “horse”?

Using the multiplication principle: 5 choices for the first letter, 4 for the second letter, 3 for the third, 2 for the fourth, and 1 for the fifth or 5!

Permutations with indistinct elements

How about for the string "hill"?

If the l's were distinct, say l_1 and l_2, it would be 4!

If the l's were distinct it would be whatever the correct answer is when they aren't distinct, x, multiplied by the possible rearrangements of 2 distinct l's, i.e. 2! x

So 4! = 2! x, or x = 4!/2!

How many permutations of the string "Illinois"?

$$\frac{8!}{3!2!} = 3360$$

Coin flips – combinations

- If I flip a coin N times, how many outcomes have exactly k heads?
- Think of this as a string of (N-k) Ts and and k Hs that is N long
- Every re-arrangement of such a string is a valid run of this experiment
- The number of such re-arrangements is "N choose k"

$$\binom{N}{k} = \frac{N!}{k!(N-k)!}$$

Probability

Counting and probability

In many cases, we will use counting in order to compute probabilities

If we are interested in an event $\mathcal{E}$, a set of outcomes, we might think of writing its probability as follows

$$P(\mathcal{E}) = \sum_{O \in \mathcal{E}} P(O)$$

Thus, if each outcome in Ω has the same probability then we have

$$P(\mathcal{E}) = \frac{\text{Number of outcomes in } \mathcal{E}}{\text{Total number of outcomes in } \Omega}$$

Example

- We stop three random people on the street and ask them on what day of the week they were born. What is the probability that they were born on days of the week that are in succession? E.g. Monday-Tuesday-Wednesday or Saturday-Sunday-Monday
- The sample space has 7^3 outcomes each equally likely
- Our event has 7 possible outcomes
- So our probability is $7/7^3 = 1/49$

Example

- Suppose we stop 2 people on the street and ask which day of the week they were born on, what is the probability they were born on the same day of the week?
- Size of the sample space is 7^2 each equally likely
- Number of outcomes in the event is 7
- Probability of our event is 1/7

Example

- A couple decides they will have three children. Let B_i be the event that there are i boys and let C be the event that there are more girls than boys.
- What is P(B_1)?
- There are 8 possible outcomes each equally likely and three of them have a single boy, so P(B_1) = 3/8
- What is P(C)?
- We could write out the outcomes and count them or we could observe that P(C) and P(C^C) are equally likely giving P(C)=1/2

Sets and probability

Another way to compute probabilities is to reason about the underlying sets of events

In particular, our rule $P(A^c) = 1 - P(A)$ along with others will be useful

Birthday problem

- If there are 30 people in a room, what is the probability that two of them celebrate their birthday on the same day?
- Let's assume that there are no leap years and that each day of the year is equally likely to be a birthday
- P({shared birthday}) = 1 - P({all 30 birthdays are different})

Birthdays

- So we will compute the probability that all 30 birthdays are different

$$P(\{\text{all birthdays different}\}) = \frac{\text{Number of outcomes in the event}}{\text{Total number of outcomes}}$$

- What is the total number of outcomes?
- 365^30 — the total number of ways to list 30 days of the year
- For our event of interest, we have 365 choices for the first birthday, 364 for the next, 363 for the next and so on

Birthdays

- So our final probability is given by

$$P(\{\text{shared birthday}\}) = 1 - \frac{365 \times 364 \times \ldots 336}{365^{30}} = 1 - 0.2937 = 0.7063$$

- Is this the same as the probability that someone in a room with 30 people has the same birthday as me?

Birthdays

- If I bet a friend that someone has the same birthday as me in a room of 30 people what is my probability of winning?
- P({winning}) = 1 - P({losing})
- For the 29 other people in the room 365^29 possible outcomes if their birthdays are random
- In how many outcomes do I wind up losing?
- 364^29
- P({winning}) = 1 - P({losing}) = 1 - 364^29/365^29
- Which is around 0.0765

Example

- We roll 2 fair dice and add the number of spots. What is the probability that we get a number that is divisible by 2 but not divisible by 5?
- Let D_n be the event corresponding to the set of outcomes where the number is divisible by n
- How might we express what we want to calculate?
- P(D_2-D_5)

Example

- A rule from a few slides ago: $P(\mathcal{A} - \mathcal{B}) = P(\mathcal{A}) - P(\mathcal{A} \cap \mathcal{B})$
- So, we are looking to compute

$$P(D_2 - D_5) = P(D_2) - P(D_2 \cap D_5)$$

- Can we simplify $P(D_2 \cap D_5)$

Example

- So we want

$$P(D_2 - D_5) = P(D_2) - P(D_{10})$$

- P(D_2)?
- 1/2
- P(D_10)?
- {4 and 6, 5 and 5, 6 and 4}, so 3/36
- 18/36 - 3/36 = 5/12

Example

- Again roll two fair dice, what is the probability that the result is divisible by 2 or 5, or both?
- If we use the D_n notation again, what do we want to compute here?

Example

- We will be computing $P(D_2 \cup D_5)$
- We can use the rule

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) - P(\mathcal{A} \cap \mathcal{B})$$

- To get

$$P(D_2 \cup D_5) = P(D_2) + P(D_5) - P(D_2 \cap D_5)$$

Example

- P(D_2) = 18/36
- P(D_5)?
- 7/36
- P(D_10) = 3/36
- 18/36 + 7/36 - 3/36 = 22/36

Independence

Independence

- Events can be related or unrelated
- Flipping a coin twice, the result of the first has no effect on what the result of the next one might be
- Informally, we call events that are unrelated **independent events**

Independence

- On the other hand, we can have events that in some way depend on one another
- If I throw a die and let A be the event that the die comes up with an odd number
- Let B be the event that the number of spots is 3 or 5
- If B has occurred, then A has occurred

Independence

- Let C be the event that the die comes up with an odd number
- Let D be the event that the number that comes up is greater than 3
- Each of these has a probability of 1/2
- However if I know that C has happened, then the die came up with 1, 3, or 5 and only one of these is in the event D
- This means knowing whether C occurred or not might tell us something about the probability that D has occurred

Independence

- We want to formalize this notion and use it, so let's make up a definition

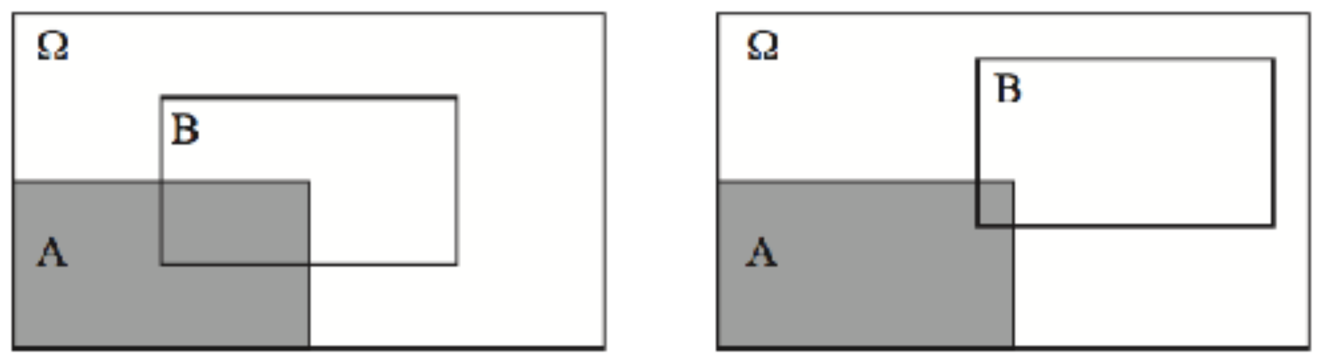
Useful Facts: 4.4 *Independent events*

Two events $\mathcal{A}$ and $\mathcal{B}$ are **independent** if and only if

$$P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A})P(\mathcal{B})$$

Independence

- Think of this definition in terms of set "sizes"



Independent Events Dependent Events

So for $\mathcal{A}$ and $\mathcal{B}$ to be independent, we must have

$$\text{“Size” of } \mathcal{A} = \frac{\text{“Size” of piece of } \mathcal{A} \text{ in } \mathcal{B}}{\text{“Size” of } \mathcal{B}},$$

Example

- If we roll two fair dice, it's reasonable to assume that what we see on one die is independent from what we see on the other
- If we want to know the probability of the event that both dice come up 3
- Let A be the event that the first is a 3 and B be the event that the second is a 3
- Using independence we get $P(A \cap B) = P(A)P(B)$
- Or that the probability of our event is $(1/6)(1/6) = 1/36$

Example

- In the example with 3 cards face down: a king, a queen, and a knight, if we draw a card, then turn it over, rearrange and draw another, what is the probability of getting two queens?
- Let A be the event that the first draw is a queen, and B be the event that the second draw is as well
- Logically, we can treat these events as independent and use $P(A \cap B) = P(A)P(B)$ to get a probability of (1/3)(1/3)=1/9 for our event

Example

- If we suppose that genders occur independently and with equal probability, what is the probability that a couple who has three children have three girls?
- $(1/2)(1/2)(1/2)=1/8$

Testing for independence

- So far we have used independence as an assumption about events
- We might also use the definition to test whether two events are in fact independent

Example

- If we draw from a shuffled standard deck of 52 cards and call A the event that the card is red and B the event that the card is a 10, **are A and B independent**?
- $P(A) = 1/2$ and $P(B) = 1/13$
- $P(AB) = 2/52$
- So $P(AB) = P(A)P(B)$, hence these events are independent

Example

- Suppose we remove the 10 of hearts from the deck, and let event C be that we draw a red card, and D be that we draw a 10. **Are C and D independent**?
- What is P(C)?
- 25/51
- What is P(D)?
- 3/51
- What is P(CD)?
- 1/51
- P(C)P(D) = 75/(51)^2
- These events aren't independent since P(C)P(D) does not equal P(CD)