Math 3081 (Probability and Statistics)

Lecture #3 of 27 $\sim$ July 8th, 2021

Probability

- Probability Distributions
- Computing Probabilities (Part 1)
- Conditional Probability

This material represents §1.3.2-1.3.4 from the course notes.

Recall

Recall some terminology from last time:

"Definition"

For any experiment or observation, the set of possible outcomes is called the sample space, and an event is a subset of the sample space.

We also decided that, if probabilities of events behave according to our intuition, they should have these properties:

1. The probability of any event should represent the proportion of times it occurs. In particular, it should always be between 0 and 1 (inclusive).
2. The probability of the entire sample space S should equal 1.
3. If E_1 and E_2 are mutually exclusive events, then $P(E_1 \cup E_2)$ should equal $P(E_1) + P(E_2)$.

Probability Distributions, I

We now use these properties to give a formal definition of a probability distribution on a sample space:

Definition

If S is a sample space, a probability distribution on S is any function P defined on events (i.e., subsets of S) with the following properties:

[P1] *For any event E, the probability $P(E)$ satisfies* $0 \leq P(E) \leq 1$.

[P2] *The probability* $P(S) = 1$.

[P3] *If $E_1, E_2, \ldots, E_k, \ldots$ are mutually exclusive events (meaning that $E_i \cap E_j = \emptyset$ for $i \neq j$), then* $P(E_1 \cup E_2 \cup \cdots \cup E_k \cup \cdots) = P(E_1) + P(E_2) + \cdots + P(E_k) + \cdots$.

Probability Distributions, II

Although the definition we have is very formal, you should keep in mind the intuitive formulation, in which the probability $P(E)$ represents the relative frequency that the event E occurs if we repeat the experiment a large number of times.

It would be helpful to give an example at this point. However, it is not so obvious how to verify condition [P3] other than by writing down all possible combinations of mutually exclusive events, which would be very time-consuming. We need to do a little more work first!

Probability Distributions, II

Consider the case of a finite sample space $S = \{s_1, s_2, \ldots, s_k\}$:

- By [P2] and [P3], $P(\{s_1\}) + P(\{s_2\}) + \cdots + P(\{s_k\}) = P(\{s_1, s_2, \ldots, s_k\}) = P(S) = 1$.
- By [P1], the numbers $P(\{s_1\})$, ... , $P(\{s_k\})$ are all between 0 and 1 inclusive, and have sum 1.
- If we make any selection for these values satisfying these conditions, then we may compute the probability of an arbitrary event $E = \{t_1, \ldots, t_d\}$ by using property [P3] again: $P(E) = P(\{t_1\}) + \cdots + P(\{t_d\})$.
- It is not hard to check that this assignment of $P(E)$ for each subset E of S satisfies all three of [P1]-[P3].

Probability Distributions, II

Consider the case of a finite sample space $S = \{s_1, s_2, \dots, s_k\}$:

- By [P2] and [P3], $P(\{s_1\}) + P(\{s_2\}) + \cdots + P(\{s_k\}) = P(\{s_1, s_2, \dots, s_k\}) = P(S) = 1$.
- By [P1], the numbers $P(\{s_1\})$, ... , $P(\{s_k\})$ are all between 0 and 1 inclusive, and have sum 1.
- If we make any selection for these values satisfying these conditions, then we may compute the probability of an arbitrary event $E = \{t_1, \dots, t_d\}$ by using property [P3] again: $P(E) = P(\{t_1\}) + \cdots + P(\{t_d\})$.
- It is not hard to check that this assignment of $P(E)$ for each subset E of S satisfies all three of [P1]-[P3].

In summary, a probability distribution on S is an assignment of probabilities between 0 and 1 inclusive to each individual outcome in the sample space, such that the sum of all the probabilities is 1.

Probability Distributions, III

Example: Consider the sample space $S = \{H, T\}$ corresponding to flipping a coin.

- A probability distribution on S is determined by assigning values to $P(H)$ and $P(T)$ such that $0 \leq P(H), P(T) \leq 1$ and $P(H) + P(T) = 1$.
- One probability distribution arises from assuming that a head is equally likely to appear as a tail: then we would have $P(H) = P(T) = 1/2$.
- A different probability distribution arises from assuming that a tail is twice as likely to appear as a head: then we would have $P(H) = 1/3$ while $P(T) = 2/3$.
- Another probability distribution, for an even more unfair coin, would have $P(H) = 0.1$ with $P(T) = 0.9$: this represents a coin that lands heads 10% of the time and tails the other 90% of the time.

Probability Distributions, IV

In the coin example, the choice of probability distribution for the sample space affects our interpretation of how fair the coin is.

- We view the probability of an event on a sliding scale from 0 to 1: a probability near 0 means that the event happens rarely, while a probability near 1 means the event happens often.
- If the sample space is finite, then an event with probability 0 never occurs, while an event with probability 1 always occurs.
- Indeed, such a probability distribution could show up in our coin setting, since the probability distribution with $P(H) = 0$ and $P(T) = 1$ is perfectly valid. (This corresponds to a coin that always lands tails.)

Properties of Probability Distributions, I

Before going on, we will establish a few other basic properties of probabilities using the definition:

Proposition (Basic Properties of Probability)

For any events E, E_1, E_2 inside a sample space S with probability distribution P, the following properties hold:

1. $P(\emptyset) = 0$.
2. $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.
3. $P(E^c) = 1 - P(E)$.

Proofs:

- For (1), apply [P3] to $E_1 = E_2 = \emptyset$, which is valid because $E_1 \cap E_2 = \emptyset$.
- Since $E_1 \cup E_2 = \emptyset$ also, this means $P(\emptyset) = P(\emptyset) + P(\emptyset)$, so $P(\emptyset) = 0$.

Properties of Probability Distributions, II

2. $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.

Proof:

- Define the events $A = E_1 \cap E_2^c$, $B = E_1 \cap E_2$, $C = E_1^c \cap E_2$.
- Now observe (since these are just the three pieces in a Venn diagram!) that A, B, C are mutually disjoint with $A \cup B = E_1$, $B \cup C = E_2$, and $A \cup B \cup C = E_1 \cup E_2$.
- Thus by [P3] applied repeatedly, we have $P(E_1 \cup E_2) = P(A \cup B \cup C) = P(A) + P(B) + P(C)$, while $P(E_1) = P(A) + P(B)$, $P(E_2) = P(B) + P(C)$, and $P(E_1 \cap E_2) = P(B)$.
- Hence we see $P(E_1 \cup E_2) = P(A) + P(B) + P(C) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$, as claimed.

Properties of Probability Distributions, III

3. $P(E^c) = 1 - P(E)$.

Proof:

- Recall that we just proved (2):
  $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.
- Now apply the result of (2) with $E_1 = E$ and $E_2 = E^c$.
- Since $E \cup E^c = S$ and $E \cap E^c = \emptyset$, applying (2) reduces to $1 = P(S) = P(E) + P(E^c) - P(\emptyset)$.
- Thus, since $P(\emptyset) = 0$ by (1) and $P(S) = 1$ by [P2], we see that $P(E^c) = 1 - P(E)$ as claimed.

Computing Probabilities, I

So far, our discussion of probability has been relatively abstract, since we have primarily spoken about probability distributions.

- In order to compute probabilities of particular events, we must make assumptions about the corresponding probability distributions, such as assuming that a coin is fair (i.e., that heads and tails are equally likely).
- In the particular case where all of the outcomes $s_1, s_2, \ldots, s_k$ in the sample space are equally likely, we would have $P(s_1) = P(s_2) = \cdots = P(s_k) = 1/k$.
- Then the probability of any event E is then simply $\#(E)/k$, which only depends on the number of outcomes in E.
- This means we can find $P(E)$ simply by counting all of the outcomes in E; in particular, when the sample space is small, we can simply list them all.

Computing Probabilities, II

Example: If a fair coin is flipped twice, determine the respective probabilities of obtaining (i) no heads, (ii) 1 head, (iii) 2 heads, and (iv) 3 heads.

Computing Probabilities, II

Example: If a fair coin is flipped twice, determine the respective probabilities of obtaining (i) no heads, (ii) 1 head, (iii) 2 heads, and (iv) 3 heads.

- The sample space has 4 outcomes: $S = \{TT, TH, HT, HH\}$.
- Under the assumption that the coin is fair, each of the 4 outcomes is equally likely, so by our analysis above, the probability of each of the 4 outcomes is $1/4$.
- There is only 1 outcome with no heads (namely, TT), so the probability of obtaining no heads is $1/4$.
- There are 2 outcomes (namely TH and HT) with 1 head, so the probability of obtaining 1 head is $2/4 = 1/2$.
- There is 1 outcome (namely, HH) with 2 heads, so the probability of obtaining 2 heads is $1/4$.
- There are no outcomes with 3 heads, so the probability of obtaining 3 heads is 0.

Computing Probabilities, III

Example: If two fair 6-sided dice are rolled, determine the probabilities of the respective events

1. The two dice read 6 and 3 in some order.
2. The sum of the two rolls is equal to 4.
3. The two rolls are equal.
4. Neither roll is a 2.
5. At least one roll is a 2.

Computing Probabilities, III

Example: If two fair 6-sided dice are rolled, determine the probabilities of the respective events

1. The two dice read 6 and 3 in some order.
2. The sum of the two rolls is equal to 4.
3. The two rolls are equal.
4. Neither roll is a 2.
5. At least one roll is a 2.

- In this case, the sample space consists of $6^2 = 36$ outcomes representing the 36 ordered pairs (R_1, R_2) where R_1 is the first roll and R_2 is the second roll.
- Under the assumption that both dice are fair, all 36 outcomes are equally likely, so the probability of each is $1/36$.
- Now we simply count to determine the probabilities.

Computing Probabilities, IV

Example: If two fair 6-sided dice are rolled, determine the probabilities of the respective events

1. The two dice read 6 and 3 in some order.

Computing Probabilities, IV

Example: If two fair 6-sided dice are rolled, determine the probabilities of the respective events

1. The two dice read 6 and 3 in some order.
   - There are two possible outcomes: $(3, 6)$ and $(6, 3)$, so the probability is $2/36 = 1/18$.
2. The sum of the two rolls is equal to 4.

Computing Probabilities, IV

Example: If two fair 6-sided dice are rolled, determine the probabilities of the respective events

1. The two dice read 6 and 3 in some order.
   - There are two possible outcomes: $(3, 6)$ and $(6, 3)$, so the probability is $2/36 = 1/18$.
2. The sum of the two rolls is equal to 4.
   - There are three possible outcomes: $(1, 3)$, $(2, 2)$, and $(3, 1)$. Thus the probability of this event is $3/36 = 1/12$.
3. The two rolls are equal.

Computing Probabilities, IV

Example: If two fair 6-sided dice are rolled, determine the probabilities of the respective events

1. The two dice read 6 and 3 in some order.
   - There are two possible outcomes: $(3, 6)$ and $(6, 3)$, so the probability is $2/36 = 1/18$.
2. The sum of the two rolls is equal to 4.
   - There are three possible outcomes: $(1, 3)$, $(2, 2)$, and $(3, 1)$. Thus the probability of this event is $3/36 = 1/12$.
3. The two rolls are equal.
   - There are six possible outcomes: $(1, 1)$, $(2, 2)$, $(3, 3)$, $(4, 4)$, $(5, 5)$, and $(6, 6)$. Thus the probability of this event is $6/36 = 1/6$.

Computing Probabilities, V

Example: If two fair 6-sided dice are rolled, determine the probabilities of the respective events

4. Neither roll is a 2.

Computing Probabilities, V

Example: If two fair 6-sided dice are rolled, determine the probabilities of the respective events

4. Neither roll is a 2.
   - There are 25 possible outcomes, consisting of the $5 \cdot 5$ ordered pairs (a, b) where each of a and b is one of the 5 numbers $1, 3, 4, 5, 6$.
   - Thus the probability of this event is $25/36$.
5. At least one roll is a 2.

Computing Probabilities, V

Example: If two fair 6-sided dice are rolled, determine the probabilities of the respective events

4. Neither roll is a 2.
   - There are 25 possible outcomes, consisting of the $5 \cdot 5$ ordered pairs (a, b) where each of a and b is one of the 5 numbers $1, 3, 4, 5, 6$.
   - Thus the probability of this event is $25/36$.
5. At least one roll is a 2.
   - One way is to write down all 11 possible outcomes explicitly, to see that the probability of this event is $11/36$.
   - Another method is to observe that this event is the complement of the event above, so its probability must be $1 - 25/36 = 11/36$.

Computing Probabilities, VI

Example: Six fair coins are flipped. Determine the probabilities of the respective events

1. The first and third flips are heads.
2. Exactly three tails are flipped.
3. The first flip and last flip are the same.
4. There are more heads flipped than tails.

Computing Probabilities, VI

Example: Six fair coins are flipped. Determine the probabilities of the respective events

1. The first and third flips are heads.
2. Exactly three tails are flipped.
3. The first flip and last flip are the same.
4. There are more heads flipped than tails.

- In this case, the sample space consists of $2^6 = 64$ outcomes representing the 64 ordered tuples $(F_1, F_2, F_3, F_4, F_5, F_6)$ where F_i is the result of the ith flip.
- Under the assumption that the coins are fair, all 64 outcomes are equally likely, so the probability of each is $1/64$.
- Now we simply count to determine the probabilities.

Computing Probabilities, VII

Example: Six fair coins are flipped. Determine the probabilities of the respective events

1. The first and third flips are heads.

Computing Probabilities, VII

Example: Six fair coins are flipped. Determine the probabilities of the respective events

1. The first and third flips are heads.
   - Here, 2 flips are fixed and the other 4 can be either heads or tails, so there are $2^4 = 16$ possible outcomes. Thus, the probability is $16/64 = 1/4$.
2. Exactly three tails are flipped.

Computing Probabilities, VII

Example: Six fair coins are flipped. Determine the probabilities of the respective events

1. The first and third flips are heads.

- Here, 2 flips are fixed and the other 4 can be either heads or tails, so there are $2^4 = 16$ possible outcomes. Thus, the probability is $16/64 = 1/4$.

2. Exactly three tails are flipped.

- There are $\binom{6}{3} = 20$ ways to select the locations of the three tails among the six flips, so there are 20 possible outcomes. Thus, the probability of this event is $20/64 = 5/16$.

Computing Probabilities, VIII

Example: Six fair coins are flipped. Determine the probabilities of the respective events

3. The first flip and last flip are the same.

Computing Probabilities, VIII

Example: Six fair coins are flipped. Determine the probabilities of the respective events

3. The first flip and last flip are the same.

- There are 2 choices for the first flip, and once this is selected the last flip is determined. Then the other 4 flips can be chosen arbitrarily. Thus there are $2^5 = 32$ possible outcomes, so the probability is $32/64 = 1/2$.

4. There are more heads flipped than tails.

Computing Probabilities, VIII

Example: Six fair coins are flipped. Determine the probabilities of the respective events

3. The first flip and last flip are the same.

- There are 2 choices for the first flip, and once this is selected the last flip is determined. Then the other 4 flips can be chosen arbitrarily. Thus there are $2^5 = 32$ possible outcomes, so the probability is $32/64 = 1/2$.

4. There are more heads flipped than tails.

- To obtain more heads than tails requires flipping 4 heads (and 2 tails), or 5 heads (and 1 tail), or 6 heads.
- There are $\binom{6}{4} = 15$, $\binom{6}{5} = 6$, and $\binom{6}{6} = 1$ ways of achieving these outcomes, respectively.
- Thus, the total number of outcomes is $15 + 6 + 1 = 22$, so the probability is $22/64 = 11/32$.

Computing Probabilities, IX

Example: Three cards are randomly dealt from a standard 52-card deck. Determine the probabilities of the respective events

1. All three cards are queens.
2. The first card is a diamond and the other two are spades.
3. One of the cards is the three of diamonds.
4. The first card is a four and the second card is a nine.
5. At least one card is a ten.

Computing Probabilities, IX

Example: Three cards are randomly dealt from a standard 52-card deck. Determine the probabilities of the respective events

1. All three cards are queens.
2. The first card is a diamond and the other two are spades.
3. One of the cards is the three of diamonds.
4. The first card is a four and the second card is a nine.
5. At least one card is a ten.

- Let us take the sample space S to be the set of possible triples (C_1, C_2, C_3) of the three cards dealt in order from the deck. By the multiplication principle, there are $52 \cdot 51 \cdot 50$ such triples, so $\#S = 52 \cdot 51 \cdot 50 = 132600$.
- Under the assumption that the cards are drawn randomly, each of the outcomes then has probability $1/132600$.

Computing Probabilities, X

<u>Example</u>: Three cards are randomly dealt from a standard 52-card deck. Determine the probabilities of the respective events

1. All three cards are queens.

Computing Probabilities, X

Example: Three cards are randomly dealt from a standard 52-card deck. Determine the probabilities of the respective events

1. All three cards are queens.

- There are 4 choices for the first card (any of the 4 queens), 3 choices for the second card (any of the 3 remaining ones), and 2 choices for the third.
- Thus the total number of outcomes is $4 \cdot 3 \cdot 2 = 24$, so the probability is $24/132600 = 1/5525$.

2. The first card is a diamond and the other two are spades.

Computing Probabilities, X

Example: Three cards are randomly dealt from a standard 52-card deck. Determine the probabilities of the respective events

1. All three cards are queens.

- There are 4 choices for the first card (any of the 4 queens), 3 choices for the second card (any of the 3 remaining ones), and 2 choices for the third.
- Thus the total number of outcomes is $4 \cdot 3 \cdot 2 = 24$, so the probability is $24/132600 = 1/5525$.

2. The first card is a diamond and the other two are spades.

- There are 13 choices for the first card (any diamond), 13 for the second card (any spade), and then 12 for the third card (any spade not already chosen).
- Thus, the total number of outcomes is $13 \cdot 13 \cdot 12 = 2028$, so the probability is $2028/132600 = 13/850 \approx 0.0153$.

Computing Probabilities, XI

Example: Three cards are randomly dealt from a standard 52-card deck. Determine the probabilities of the respective events

3. One of the cards is the three of diamonds.

Computing Probabilities, XI

Example: Three cards are randomly dealt from a standard 52-card deck. Determine the probabilities of the respective events

3. One of the cards is the three of diamonds.

- There are 3 choices for which card is the 3 of diamonds. Once this is chosen, there are 51 possibilities for the first remaining card (any card except the 3 of diamonds) and 50 possibilities for the second remaining card.
- This gives a total of $3 \cdot 51 \cdot 50 = 7650$ outcomes, so the probability is $7650/132600 = 3/52 \approx 0.0577$.
- Intuitively, since we are choosing 3 cards out of 52, and the 3 of diamonds is equally likely to be any one of these 52 cards, the probability that it is among the 3 cards we have chosen should be $3/52$.

Computing Probabilities, XII

Example: Three cards are randomly dealt from a standard 52-card deck. Determine the probabilities of the respective events

4. The first card is a four and the second card is a nine.

Computing Probabilities, XII

Example: Three cards are randomly dealt from a standard 52-card deck. Determine the probabilities of the respective events

4. The first card is a four and the second card is a nine.
   - There are 4 choices for the first card (any four), 4 for the second card (any nine), and then 50 for the third card (any card not already chosen).
   - Thus, the total number of outcomes is $4 \cdot 4 \cdot 50 = 800$, so the probability is $800/132600 = 4/663 \approx 0.0060$.
5. At least one card is a ten.

Computing Probabilities, XII

Example: Three cards are randomly dealt from a standard 52-card deck. Determine the probabilities of the respective events

4. The first card is a four and the second card is a nine.
   - There are 4 choices for the first card (any four), 4 for the second card (any nine), and then 50 for the third card (any card not already chosen).
   - Thus, the total number of outcomes is $4 \cdot 4 \cdot 50 = 800$, so the probability is $800/132600 = 4/663 \approx 0.0060$.
5. At least one card is a ten.
   - We instead count the complement: in that case there are 48 choices for the first card (anything but the 4 tens), 47 for the second, and 46 for the third, totaling $48 \cdot 47 \cdot 46 = 103776$.
   - Thus, the total number of outcomes with at least one ten is $132600 - 103776 = 28824$, so the probability is $28824/132600 = 1201/5525 \approx 0.2174$.

Computing Probabilities, XIII

<u>Example</u>: A fair coin is flipped 14 times. Determine the probabilities of the respective events

1. The first five flips are all heads.
2. Exactly 5 tails are obtained.
3. Fewer than 3 tails are obtained.
4. At least 2 heads are obtained.
5. The number of tails flipped is even.

Computing Probabilities, XIII

Example: A fair coin is flipped 14 times. Determine the probabilities of the respective events

1. The first five flips are all heads.
2. Exactly 5 tails are obtained.
3. Fewer than 3 tails are obtained.
4. At least 2 heads are obtained.
5. The number of tails flipped is even.

- In this case, the sample space consists of $2^{14} = 16384$ outcomes representing the ordered tuples $(F_1, F_2, \ldots, F_14)$ where F_i is the result of the ith flip.
- Under the assumption that the coins are fair, each event has probability $1/16384$.

Computing Probabilities, XIV

Example: A fair coin is flipped 14 times. Determine the probabilities of the respective events

1. The first five flips are all heads.

Computing Probabilities, XIV

Example: A fair coin is flipped 14 times. Determine the probabilities of the respective events

1. The first five flips are all heads.

- Here, 5 flips are fixed and the other 9 can be either heads or tails, so there are 2^9 possible outcomes. Thus, the probability is $2^9/2^{14} = 1/32$.

2. Exactly 5 tails are obtained.

Computing Probabilities, XIV

Example: A fair coin is flipped 14 times. Determine the probabilities of the respective events

1. The first five flips are all heads.

- Here, 5 flips are fixed and the other 9 can be either heads or tails, so there are 2^9 possible outcomes. Thus, the probability is $2^9/2^{14} = 1/32$.

2. Exactly 5 tails are obtained.

- There are $\binom{14}{5} = 2002$ ways to select the locations of the five tails among the 14 flips, so there are 2002 possible outcomes. Thus, the probability of this event is $2002/2^{14} = 1001/8192 \approx 0.1222$.

Computing Probabilities, XV

<u>Example</u>: A fair coin is flipped 14 times. Determine the probabilities of the respective events

3. Fewer than 3 tails are obtained.

Computing Probabilities, XV

Example: A fair coin is flipped 14 times. Determine the probabilities of the respective events

3. Fewer than 3 tails are obtained.

   - There are $\binom{14}{k}$ ways to obtain k tails, so summing over the possibilities of 0, 1, and 2 tails yields a total of $\binom{14}{0} + \binom{14}{1} + \binom{14}{2} = 1 + 14 + 91 = 106$ outcomes.
   - Thus, the probability is $106/2^{14} = 53/8192 \approx 0.0065$.

4. At least 2 heads are obtained.

Computing Probabilities, XV

Example: A fair coin is flipped 14 times. Determine the probabilities of the respective events

3. Fewer than 3 tails are obtained.

- There are $\binom{14}{k}$ ways to obtain k tails, so summing over the possibilities of 0, 1, and 2 tails yields a total of $\binom{14}{0} + \binom{14}{1} + \binom{14}{2} = 1 + 14 + 91 = 106$ outcomes.
- Thus, the probability is $106/2^{14} = 53/8192 \approx 0.0065$.

4. At least 2 heads are obtained.

- In this case it is easier to count the complement ("fewer than 2 heads are obtained"), which has $\binom{14}{0} + \binom{14}{1} = 15$ possible outcomes.
- Thus, the probability of the complement is $15/2^{14}$ so the probability of the original event is $1 - 15/2^{14}$.

Computing Probabilities, XVI

Example: A fair coin is flipped 14 times. Determine the probabilities of the respective events

5. The number of tails flipped is even.

Computing Probabilities, XVI

Example: A fair coin is flipped 14 times. Determine the probabilities of the respective events

5. The number of tails flipped is even.

- There are $\binom{14}{k}$ ways to obtain k tails, so summing over the possibilities of 0, 2, 4, ... , 14 tails yields a total of $\binom{14}{0} + \binom{14}{2} + \cdots + \binom{14}{14} = 8192$ outcomes.
- Thus, the probability is $8192/2^{14} = 1/2$.

Computing Probabilities, XVI

<u>Example</u>: A fair coin is flipped 14 times. Determine the probabilities of the respective events

5. The number of tails flipped is even.

- There are $\binom{14}{k}$ ways to obtain k tails, so summing over the possibilities of 0, 2, 4, ... , 14 tails yields a total of $\binom{14}{0} + \binom{14}{2} + \cdots + \binom{14}{14} = 8192$ outcomes.
- Thus, the probability is $8192/2^{14} = 1/2$.

Notice that this probability is exactly $1/2$. In fact, there is a slick argument to explain why:

- Imagine flipping the first 13 coins. To obtain an even number of tails, there is then exactly one possible outcome for the remaining coin.
- Thus, there are exactly 2^{13} possible outcomes with an even number of tails, so the probability is $1/2$.

Conditional Probability: Motivation, I

So far, we have calculated probabilities under the assumption of "fairness", where all outcomes in the sample space are equally likely. But our use of probability distributions is flexible enough to handle general situations where outcomes are *not* equally likely.

- However, in such a case, we must also assign a probability to each individual outcome in the events we are studying.
- In the event of simple situations like flipping an unfair coin once or rolling a loaded die once, this is not so difficult, since there are very few outcomes in the sample space.
- But in more complicated scenarios (e.g., flipping an unfair coin 10 times), it is less clear how to assign probabilities.
- Therefore, we will first discuss conditional probability, which provides us a way to compute the probability that one event occurs, given that another event also occurred.

Conditional Probability; Motivation, II

Example: Two fair 6-sided dice are rolled. Given that the sum of the two rolls is 5, find the probability that the first roll was a 4.

Conditional Probability; Motivation, II

Example: Two fair 6-sided dice are rolled. Given that the sum of the two rolls is 5, find the probability that the first roll was a 4.

- First, notice that of the 36 possible outcomes from rolling the two dice, there are only 4 outcomes that are consistent with the given information: namely, $(4,1)$, $(3,2)$, $(2,3)$, and $(1,4)$.
- Since each of these outcomes was equally likely to occur originally, it is reasonable to say that they should still be equally likely.
- Our desired event (where the first roll was a 4) occurs in 1 of these 4 outcomes, so the probability of the event is $1/4$.

In this example, notice that by providing additional information (namely, that the sum of the two rolls is 5), the probability that the first roll was a 4 changes from $1/6$ to $1/4$. This is the essential idea of conditional probability.

Conditional Probability, Motivation; II

Example: Suppose that 100 students in a course have grades and academic standing as given in the table below.

Category	A	B	C	Total
Sophomore	4	8	6	18
Junior	14	11	10	35
Senior	38	4	5	47
Total	56	23	21	100

Compute the probabilities that

1. A randomly-chosen student is getting an A.
2. A randomly-chosen student is a junior.
3. A randomly-chosen junior is getting a A.
4. A randomly-chosen A student is a junior.

Conditional Probability? Motivation! IV

Category	A	B	C	Total
Sophomore	4	8	6	18
Junior	14	11	10	35
Senior	38	4	5	47
Total	56	23	21	100

1. A randomly-chosen student is getting an A.

- There are 56 A-students out of 100 total students, so the probability is $56/100 = 0.56$.

2. A randomly-chosen student is a junior.

- There are 35 juniors out of 100 total students, so the probability is $35/100 = 0.35$.

Conditional Probability! Motivation – V

Category	A	B	C	Total
Sophomore	4	8	6	18
Junior	14	11	10	35
Senior	38	4	5	47
Total	56	23	21	100

3. A randomly-chosen junior is getting a A.

- There are 14 A-student juniors out of 35 total juniors, so the probability is $14/35 = 0.4$.

4. A randomly-chosen A student is a junior.

- There are 14 A-student juniors out of 56 total A-students, so the probability is $14/56 = 0.25$.

Conditional, Probability. Motivation? VI!

Of the four probabilities we just calculated:

1. A randomly-chosen student is getting an A: 0.56.
2. A randomly-chosen student is a junior: 0.35.
3. A randomly-chosen junior is getting a A: 0.4.
4. A randomly-chosen A student is a junior: 0.25.

notice that event 3 is an example of a conditional probability, namely, the probability that a student is getting an A given that the student is a junior.

Conditional, Probability. Motivation? VI!

Of the four probabilities we just calculated:

1. A randomly-chosen student is getting an A: 0.56.
2. A randomly-chosen student is a junior: 0.35.
3. A randomly-chosen junior is getting a A: 0.4.
4. A randomly-chosen A student is a junior: 0.25.

notice that event 3 is an example of a conditional probability, namely, the probability that a student is getting an A given that the student is a junior.

In order to compute this conditional probability, notice that we restricted our attention only to the set of juniors, and performed our computations as if this were our entire sample space.

Event 4 is also a conditional probability (think about why).

Conditional% Probability@ Motivation/ VII¿

Let's examine this a bit more closely:

- Let S be the full sample space of all 100 students, let J represent the event that a student is a junior and let A represent the event that a student receives an A.
- If we write $P(A|J)$ for the probability that a student is receiving an A given that they are a junior, which is event 3.
- Then $P(A|J)$ is the ratio of juniors receiving an A to the total number of juniors: symbolically, $P(A|J) = \dfrac{\#(J \cap A)}{\#(J)}$.
- Now observe that $P(A|J) = \dfrac{\#(J \cap A)}{\#(J)} = \dfrac{\#(J \cap A)/\#(S)}{\#(J)/\#(S)} = \dfrac{P(J \cap A)}{P(J)}$.
- This last observation gives us a way to write $P(A|J)$ in terms of probabilities of events in the original sample space.

Conditional Probability, I

We can now give the general definition of conditional probabilities:

Definition

If A and B are events and $P(B) > 0$, we define the <u>conditional probability</u> $P(A|B)$, the probability that A occurs given that B occurred, as $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$.

- If $P(B) = 0$, then the conditional probability $P(A|B)$ cannot be computed using this definition.
- When the sample space is finite, then the conditional probability $P(A|B)$ does not make sense if $P(B) = 0$, because the event B can never occur.

Conditional Probability, II

Example: Suppose two fair 6-sided dice are rolled. If A is the event "neither roll is a 2" and B is the event "the first roll is a 5", describe what $P(A|B)$ and $P(B|A)$ mean.

Conditional Probability, II

Example: Suppose two fair 6-sided dice are rolled. If A is the event "neither roll is a 2" and B is the event "the first roll is a 5", describe what $P(A|B)$ and $P(B|A)$ mean.

- By definition, $P(A|B)$ is the probability that A occurs given that B occurred, so in this case it is the probability that neither roll is a 2, given that the first roll is a 5.
- Inversely, $P(B|A)$ is the probability that B occurs given that A occurred, so in this case it is the probability that the first roll is a 5 given that neither roll is a 2.

Conditional Probability, III

Example: Suppose two fair 6-sided dice are rolled. If A is the event "neither roll is a 2" and B is the event "the first roll is a 5", calculate $P(A|B)$ and $P(B|A)$.

Conditional Probability, III

Example: Suppose two fair 6-sided dice are rolled. If A is the event "neither roll is a 2" and B is the event "the first roll is a 5", calculate $P(A|B)$ and $P(B|A)$.

- We have $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$ and $P(B|A) = \dfrac{P(A \cap B)}{P(A)}$ by the definition of these conditional probabilities, so we need only calculate $P(A)$, $P(B)$, and $P(A \cap B)$.
- By listing outcomes, we can see $P(A) = 25/36$, $P(B) = 1/6$, and $P(A \cap B) = 5/36$.
- Hence the formulas give $P(A|B) = \dfrac{5/36}{1/6} = \dfrac{5}{6}$, and also $P(B|A) = \dfrac{5/36}{25/36} = \dfrac{1}{5}$.

Summary

We discussed probability distributions from a general perspective.

We used counting techniques to compute probabilities when all outcomes in the sample space are equally likely.

We introduced conditional probability.

Next lecture: Conditional probability and independence.