Many probability questions are concerned with some numerical value associated with an experiment.

- Number of 1 bits generated
- Number of "heads" flips
- Beats per minute of a heart
- Number of boys in a family
- Longevity of a chicken
**Definition:** A random variable is a function from the sample space of an experiment to the set of real numbers \( \mathbb{R} \). That is, a random variable assigns a real number to each possible outcome.

**Example:** Suppose that a coin is flipped three times. Let \( X(s) \) be the random variable that equals the numbers of heads that appear when \( s \) is the outcome. Then \( X(s) \) takes the following values:

- \( X(\text{HHH}) = 3 \)
- \( X(\text{HHT}) = X(\text{HTH}) = X(\text{THH}) = 2 \)
- \( X(\text{TTH}) = X(\text{THT}) = X(\text{HTT}) = 1 \)
- \( X(\text{TTT}) = 0 \)

**Note:** \( X \) is not a variable, and is not random. \( X \) is a function!
Definition: The distribution of a random variable $X$ on a sample space $S$ is the set of pairs $(r, p(X=r))$ for all $r \in X(S)$, where $p(X=r)$ is the probability that $X$ takes the value $r$.

Note: A distribution is usually described by specifying $p(X=r)$ for each $r \in X(S)$.

Example: Assume that our coin flips from the previous slide were all equally likely to occur. We then get the following distribution for the random variable $X$:

- $p(X=0) = 1/8$
- $p(X=1) = 3/8$
- $p(X=2) = 3/8$
- $p(X=3) = 1/8$
Many times, we want to study the expected value of a random variable

**Definition:** The expected value (or expectation) of a random variable \( X(s) \) on the sample space \( S \) is equal to:

\[
E(X) = \sum_{s \in S} p(s)X(s)
\]

For every outcome... use the probability of that outcome occuring... to weight the value of the random variable for that outcome.

**Note:** The expected value of a random variable defined on an infinite sample space is defined iff the infinite series in the definition is absolutely convergent.
Example: Let $X$ be the number that comes up when a die is rolled. What is the expected value of $X$?

Solution:

- 6 possible outcomes: 1, 2, 3, 4, 5, 6
- Each outcome occurs with the probability $1/6$
- $E(X) = 1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 6/6$
- \[ = 21/6 \]
- \[ = 7/2 \]
**Example:** A fair coin is flipped three times. Let $S$ be the sample space of the eight possible outcomes, and $X$ be the random variable that assigns to an outcome the number of heads in that outcome. What is the expected value of $X$?

**Solution:**

- Since coin flips are independent, each outcome is equally likely
- $E(X) = 1/8[X(\text{HHH}) + X(\text{HHT}) + X(\text{HTH}) + X(\text{THH}) + X(\text{TTT}) + X(\text{TTH}) + X(\text{THT}) + X(\text{TTT})]$
- $= 1/8[3 + 2 + 2 + 2 + 1 + 1 + 1 + 0]$
- $= 12/8$
- $= 3/2$
If S is large, the definition of expected value can be difficult to use directly

**Definition:** If X is a random variable and p(X=r) is the probability that X = r (i.e., p(X=r) = \( \sum_{s \in S, X(s)=r} p(s) \)), then

\[
E(X) = \sum_{r \in X(S)} p(X = r)r
\]

Each value of X... 

... is weighted by its probability of occurrence.

**Proof:**

- Suppose that X is a random variable ranging over S
- Note that p(X=r) is the probability that X takes the value r
- This means that p(X=r) is the sum of the probabilities of the outcomes s\( \in S \) such that X(s) = r
- It thus follows that \( E(X) = \sum_{r \in X(S)} p(X = r)r \). 

\( \square \)
**Example:** Let $X$ be the sum of the numbers that appear when a pair of fair dice is rolled. What is the expected value of $X$?

**Recall from last week:**
- $X(1,1) = 2$ \( p(X=2) = 1/36 \)
- $X(1,2) = X(2,1) = 3$ \( p(X=3) = 2/36 = 1/18 \)
- $X(1,3) = X(2,2) = X(3,1) = 4$ \( p(X=4) = 3/36 = 1/12 \)
- $X(1,4) = X(2,3) = X(3,2) = X(4,1) = 5$ \( p(X=5) = 4/36 = 1/9 \)
- $X(1,5) = X(2,4) = X(3,3) = X(4,2) = X(5,1) = 6$ \( p(X=6) = 5/36 \)
- $X(1,6) = X(2,5) = X(3,4) = X(4,3) = X(5,2) = X(6,1) = 7$ \( p(X=7) = 6/36 = 1/6 \)
- $X(2,6) = X(3,5) = X(4,4) = X(5,3) = X(6,2) = 8$ \( p(X=8) = 5/36 \)
- $X(3,6) = X(4,5) = X(5,4) = X(6,3) = 9$ \( p(X=9) = 4/36 = 1/9 \)
- $X(4,6) = X(5,5) = X(6,4) = 10$ \( p(X=10) = 3/36 = 1/12 \)
- $X(5,6) = X(6,5) = 11$ \( p(X=11) = 2/36 = 1/18 \)
- $X(6,6) = 12$ \( p(X=12) = 1/36 \)

**So we have that:**
- $E(X) = 2(1/36) + 3(1/18) + 4(1/12) + 5(1/9) + 6(5/36) + 7(1/6) + 8(5/36) + 9(1/9) + 10(1/12) + 11(1/18) + 12(1/36)$
- $= 7$
We can apply this formula to reason about Bernoulli trials!

**Theorem:** The expected number of successes when $n$ independent Bernoulli trials are performed, in which $p$ is the probability of success, is $np$.

**Proof:**

- Let $X$ be a random variable equal to the number of successes in $n$ trials.
- We know from last week that $p(X=k) = C(n,k)p^kq^{n-k}$. So:

\[
E(X) = \sum_{k=1}^{n} kp(X = k)
\]

Definition of $E(X)$

\[
= \sum_{k=1}^{n} kC(n,k)p^k q^{n-k}
\]

Probability of $k$ successes in $n$ trials

\[
= \sum_{k=1}^{n} nC(n-1,k-1)p^k q^{n-k}
\]

Lemma: $kC(n,k) = nC(n-1,k-1)$
Proof (continued)

- \[ np \sum_{k=1}^{n} C(n-1, k-1)p^{k-1}q^{n-k} \quad \text{Factor out np from each term} \]

- \[ np \sum_{j=1}^{n-1} C(n-1, j)p^{j}q^{n-1-j} \quad \text{Shift index } j = k-1 \]

- \[ = np(p + q)^{n-1} \quad \text{Binomial theorem} \]

- \[ = np \quad \text{p+q = 1 \, \Box} \]

Note: As long as we can prove that \( kC(n,k) = nC(n-1,k-1) \), the theorem has been proved, since we have shown that \( np \) is the expected number of successes in \( n \) independent Bernoulli trials.
Proof of Lemma

Lemma: \( kC(n,k) = nC(n-1,k-1) \)

Proof:

- \( kC(n,k) = \frac{k(n!)}{[k!(n-k)!]} \)  
  \( \quad \text{Definition of } C(n,k) \)
- \( = \frac{n!}{[(k-1)!(n-k)!]} \)  
  \( \quad \text{Cancel out } k \text{ term} \)
- \( = \frac{n(n-1)!}{[(k-1)!(n-k)!]} \)  
  \( \quad \text{Factor out } n \text{ from numerator} \)
- \( = \frac{n(n-1)!}{[(k-1)!(n-1)-(k-1))!]} \)  
  \( \quad n-k = (n-1) - (k-1) \)
- \( = nC(n-1,k-1) \)  
  \( \quad \blacksquare \)
**Theorem:** If $X_1, X_2, ..., X_n$ are random variables on $S$ and if $a$ and $b$ are real numbers, then

1. $E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$
2. $E(aX + b) = aE(X) + b$

**Proof:**

- To prove the first result for $n=2$, note that
  \[ E(X_1 + X_2) = \sum_{s \in S} p(s)(X_1(s) + X_2(s)) \]
  \[ = \sum_{s \in S} p(s)X_1(s) + \sum_{s \in S} p(s)X_2(s) \]
  \[ = E(X_1) + E(X_2) \]

- The case with $n$ variables is an easy proof by induction

- To prove the second property, note that
  \[ E(aX + b) = \sum_{s \in S} p(s)(aX(s) + b) \]
  \[ = \sum_{s \in S} p(s)aX(s) + \sum_{s \in S} p(s)b \]
  \[ = a\sum_{s \in S} p(s)X(s) + b\sum_{s \in S} p(s) \]
  \[ = aE(X) + b \]

$\blacksquare$
Example: What is the expected value of the sum of the numbers that appear when two fair dice are rolled?

Solution:

- Let $X_1$ and $X_2$ be random variables indicating the value on the first and second die, respectively
- Want to calculate $E(X_1 + X_2)$
- By the previous theorem, we have that $E(X_1 + X_2) = E(X_1) + E(X_2)$
- From earlier in lecture, we know that $E(X_1) = E(X_2) = 7/2$
- So, $E(X_1 + X_2) = 7/2 + 7/2 = 7$

Note: This agrees with the (more complicated) calculation that we made earlier in lecture.
Observation: We can formulate many hard problems in terms of the sum of much easier problems!
The forgetful coat-check clerk

**Example:** A careless coat-check clerk takes the coats of \( n \) people at a restaurant, but forgets to attach the claim number to each coat. When customers return for their coats, the clerk simply returns a coat at random. What is the expected number of coats returned correctly?

**Solution:**

- Let \( X \) be the random variable that equals the number of people who get back the correct coat.
- Let \( X_i \) be the random variable with \( X_i = 1 \) if person \( i \) gets the correct coat back, and \( X_i = 0 \) otherwise.
- It follows that \( X = X_1 + X_2 + ... + X_n \).
- Since any coat can be returned to any person, the probability that person \( i \) gets back the right coat is \( \frac{1}{n} \).
- So \( E(X_i) = 1 \times p(X_i=1) + 0 \times p(X_i=0) = \frac{1}{n} + 0 = \frac{1}{n} \).
- By the linearity of expectations, we have that
  \[
  E(X) = E(X_1) + E(X_2) + ... + E(X_n)
  \]
  \[
  = \frac{1}{n} + \frac{1}{n} + ... + \frac{1}{n}
  \]
  \[
  = 1
  \]
**Definition:** The ordered pair \((j,k)\) is called an inversion in a permutation if \(j < k\), but \(k\) precedes \(j\) in the permutation. For example, the permutation 3,1,2 contains two inversions. Namely (1,3) and (2,3).

**Example:** What is the expected number of inversions in a permutation of the first \(n\) positive integers?

**Solution:**
- Let \(X\) be a RV equal to the number of inversions in a permutation.
- Let \(I_{j,k}\) be a RV such that \(I_{j,k} = 1\) if \((j,k)\) is an inversion, and \(I_{j,k} = 0\) otherwise.
- It follows that \(E(X) = \sum_{1 \leq j < k \leq n} E(I_{j,k})\).
- **Note:** In any permutation, it is equally likely for \(j\) to precede \(k\) as it is for \(k\) to precede \(j\).
- As a result, we have that \(E(I_{j,k}) = 1/2\) for all \(j,k\).
- Since there \(C(n,2)\) pairs \(j,k\) with \(1 \leq j < k \leq n\) we have that
  - \(E(X) = \sum_{1 \leq j < k \leq n} E(I_{j,k}) = 1/2 \times C(n,2) = n(n-1)/4\).
**Definition:** Random variables $X$ and $Y$ on a sample space $S$ are independent if $p(X(s)=r_1 \text{ and } Y(s)=r_2) = p(X(s)=r_1)p(Y(s)=r_2)$ for all real numbers $r_1$ and $r_2$.

**Example:** Let $X_1$ and $X_2$ be two random variables that take the values of two dice rolled. Are these two RVs independent?

**Solution:**
- $S = \{1, 2, 3, 4, 5, 6\}$
- Let $i, j \in S$
- Note that there are 36 equally likely possible outcomes for two dice
- So $p(X_1=i \text{ and } X_2=j) = 1/36$
- Note also that $p(X_1=i) = p(X_2=j) = 1/6$
- So $p(X_1=i)p(X_2=j) = 1/36$
- As a result, we can conclude that $X_1$ and $X_2$ are independent
Not all random variables are independent!

**Example:** Let $X_1$ and $X_2$ be random variables defined as before. Show that $X_1$ and $X = X_1 + X_2$ are not independent.

**Solution:**

- Note that $p(X_1=1 \text{ and } X=12) = 0$
  - Why? $X_1=1$ means that $X_1 + X_2 = 12$ is impossible!
- Note also that $p(X_1=1) = 1/6$
- Further, $p(X=12) = 1/36$
- This means that $p(X_1=1)p(X=12) \neq p(X_1=1 \text{ and } X=12)$
- By definition $X_1$ and $X$ are not independent
**Theorem:** If $X$ and $Y$ are independent random variables on a sample space $S$, then $E(XY) = E(X)E(Y)$.

**Note:** This theorem only works if $X$ and $Y$ are independent!

For example, let $X$ and $Y$ be two random variables that count the number of heads and tails, respectively, when a coin is flipped twice. Clearly $X$ and $Y$ are not independent.

We know that $p(X=2) = p(Y=2) = 1/4$, $p(X=1) = p(Y=1) = 1/2$, and $p(X=0) = p(Y=0) = 1/4$. So, $E(X) = E(Y) = 1$.

Now, note that $XY=0$ if either two heads or two tails are flipped and $XY=1$ when one head and one tail come up. This gives us that $E(XY) = 1(1/2) + 0(1/2) = 1/2$.

However, $E(X)E(Y) = 1$, so $E(XY) \neq E(X)E(Y)$. 
Problem 1: Consider a die in which the number 5 is two times as likely to be rolled as any other number. What is the expected value of this die?

Problem 2: Alice and Bob regularly play chess together. Historically, Alice wins 70% of the time. If Alice and Bob play 7 games of chess, how many games can Alice be expected to win?

Problem 3: A test contains 50 T/F questions, each worth two points, and 25 multiple choice questions, each worth four points. The probability that Alice answers and T/F question correctly is 0.9. The probability that Alice answers a multiple choice question correctly is 0.8. What is her expected score on the final?
Sometimes we need more information than the expected value can give us.

The expected value of a random variable doesn’t tell us the whole story...

\[ p(X(s)=r) \]

\[ X(s) \]
The variance of a random variable gives us information about how wide it is spread.

**Definition:** The variance of a random variable $X$ on a sample space $S$ is defined as:

$$V(X) = \sum_{s \in S} (X(s) - E(X))^2p(s)$$

- **Squared difference from expected value**
- **Weighted by probability of occurrence**

**Definition:** The standard deviation of a random variable $X$ on a sample space $S$ is defined as $\sqrt{V(S)}$. 
**Theorem:** If $X$ is a random variable on a sample space $S$, then $V(X) = E(X^2) - E(X)^2$.

**Proof:**

- $V(X) = \sum_{s \in S} (X(s) - E(X))^2 p(s)$
- $= \sum_{s \in S} X(s)^2 p(s) - 2E(X)\sum_{s \in S} X(s)p(s) + E(X)^2\sum_{s \in S} p(s)$
- $= E(X^2) - 2E(X)E(X) + E(X)^2$
- $= E(X^2) - E(X)^2 \quad \square$
Variance of a Bernoulli Distribution

**Example:** What is the variance of random variable $X$ with $X(t)=1$ if a Bernoulli trial is a success and $X(t)=0$ otherwise? Assume that the probability of success is $p$.

**Solution:**

- Note that $X$ takes only the values 0 and 1
- Hence, $X(t) = X^2(t)$
- $V(X) = E(X^2) - E(X)^2$
- $= p - p^2$
- $= p(1-p)$
- $= pq$

This tells us that the variance of ANY Bernoulli distribution is $pq$!
Example: Two dice are rolled. What is the variance of the random variable $X((j,k)) = 2j$, where $j$ is the number appearing on the first die and $k$ is the number appearing on the second die.

Solution:

- $V(X) = E(X^2) - E(X)^2$
- Note that $p(X=k) = 1/6$ for $k = 2, 4, 6, 8, 10, 12$ and is 0 otherwise
- $E(X) = (2+4+6+8+10+12)/6 = 7$
- $E(X^2) = (2^2+4^2+6^2+8^2+10^2+12^2) = 182/3$
- So $V(X) = 182/3 - 49 = 35/3$
Problem: Find the variance of a random variable $X$ where $X$ is the number that comes up when a single die is rolled.
The variance of independent random variables

**Theorem:** Let $X$ and $Y$ be independent random variables on a sample space $S$. Then, $V(X+Y) = V(X) + V(Y)$. More generally, if $X_1, X_2, \ldots, X_n$ are pairwise independent random variables on a sample space $S$, then $V(X_1+X_2+\ldots+X_n) = V(X_1) + V(X_2) + \ldots + V(X_n)$.

**Proof (2 variable case):**

- $V(X+Y) = E((X+Y)^2) - E(X+Y)^2$
- $= E(X^2 + 2XY + Y^2) - (E(X)+E(Y))^2$
- $= E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2$
- $= [E(X^2) - E(X)^2] + [E(Y^2)-E(Y)^2] + 2E(X)E(Y) - 2E(X)E(Y)$
- $= [E(X^2) - E(X)^2] + [E(Y^2)-E(Y)^2]$  
- $= V(X) + V(Y) \quad \Box$

Since $X$ and $Y$ are independent...
Example: The X be a random variable whose value is the sum that appears when two dice are rolled. What is V(X)?

Solution:
- Let $X_1$ and $X_2$ be random variables taking the value that appears on the first and second die, respectively
- So $X = X_1 + X_2$
- Note that $X_1$ and $X_2$ are independent
- $V(X) = V(X_1 + X_2)$
- $V(X_1) + V(X_2) = \frac{35}{12} + \frac{35}{12} = \frac{35}{6}$
Problem: What is the variance of the number of successes when \( n \) independent Bernoulli trials are performed, where \( p \) is the probability of success for each trial? Hint: Let \( X_i((t_1, ..., t_n)) \) be a random variable such that \( X_i = 1 \) if \( t_i \) was a success and \( X_i = 0 \) otherwise.
Final Thoughts

- Analyzing the expected value of a random variable allows us to answer a range of interesting questions.

- The variance of a random variable tells us about the spread of values that the random variable can take.