# Introduction to Probability: Applications, distributions, expectation

<table>
<thead>
<tr>
<th>Lecture A</th>
<th>Tiefenbruck</th>
<th>Tu, Th 11am-12:20pm</th>
<th>Center 119</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lecture B</td>
<td>Tiefenbruck</td>
<td>Tu, Th 9:30am-10:50am</td>
<td>Center 119</td>
</tr>
<tr>
<td>Lecture C</td>
<td>Minnes</td>
<td>Tu, Th 3:30pm-4:50pm</td>
<td>WLH 2005</td>
</tr>
</tbody>
</table>

[http://cseweb.ucsd.edu/classes/fa15/cse21-abc/](http://cseweb.ucsd.edu/classes/fa15/cse21-abc/)

Nov 19, 2015
Probability

Sample space, $S$: (finite or countable) set of possible outcomes.

Probability distribution, $p$: assignment of probabilities to outcomes in $S$ so that

- $0 \leq p(s) \leq 1$ for each $s$ in $S$.
- Sum of probabilities is 1, $\sum_{s \in S} p(s) = 1$. 

Rosen p. 446. 453
Sample space, $S$: (finite or countable) set of possible outcomes.

Probability distribution, $p$: assignment of probabilities to outcomes in $S$ so that

- $0 \leq p(s) \leq 1$ for each $s$ in $S$.
- Sum of probabilities is 1, $\sum_{s \in S} p(s) = 1$.

Compare flipping a fair coin and a biased coin:

A. Have different sample spaces.
B. Have the same sample spaces but different probability distributions.
C. Have the same sample space and same probability distributions.
Probability

Sample space, $S$: (finite or countable) set of possible outcomes.

Probability distribution, $p$: assignment of probabilities to outcomes in $S$ so that

- $0 \leq p(s) \leq 1$ for each $s$ in $S$.
- Sum of probabilities is 1, $\sum_{s \in S} p(s) = 1$.

Event, $E$: subset of possible outcomes. $P(E) = \sum_{s \in E} p(s)$.
For sample space S with n elements, uniform distribution assigns the probability 1/n to each element of S.

When flipping a fair coin successively three times:

A. The sample space is \{H, T\}
B. The empty set is not an event.
C. The event \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\} has probability less than 1.
D. The uniform distribution assigns probability 1/8 to each outcome.
E. None of the above.
Uniform distribution

For sample space $S$ with $n$ elements, **uniform distribution** assigns the probability $1/n$ to each element of $S$.

When flipping a fair coin successively three times:

A. The sample space is $\{H, T\}$
B. The empty set is not an event.
C. The event $\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$ has probability less than 1.
D. The uniform distribution assigns probability $1/8$ to each outcome.
E. None of the above.
For sample space S with n elements, **uniform distribution** assigns the probability 1/n to each element of S.

\[
P(0\,H) = \frac{1}{8}, \quad P(1\,H) = \frac{3}{8}, \quad P(2\,H) = \frac{3}{8}
\]

When flipping a fair coin successively three times, what is the distribution of the number of Hs that appear?

A. Uniform distribution.
B. \(P(0\,H) = P(3\,H) = \frac{3}{8}\) and \(P(1\,H) = P(2\,H) = \frac{1}{8}\).
C. \(P(0\,H) = P(1\,H) = \frac{1}{8}\) and \(P(2\,H) = P(3\,H) = \frac{1}{8}\).
D. \(P(0\,H) = P(3\,H) = \frac{1}{8}\) and \(P(1\,H) = P(2\,H) = \frac{1}{8}\).
E. None of the above.
For sample space S with n elements, **uniform distribution** assigns the probability $1/n$ to each element of S.

When flipping a fair coin successively three times, what is the **distribution of the number** of Hs that appear?

A. Uniform distribution.
B. $P(0\ H) = P(3\ H) = 3/8$ and $P(1\ H) = P(2\ H) = 1/8$.
C. $P(0\ H) = P(1\ H) = 1/8$ and $P(2\ H) = P(3\ H) = 1/8$.
D. $P(0\ H) = P(3\ H) = 1/8$ and $P(1\ H) = P(2\ H) = 1/8$.
E. None of the above.

**Not a uniform distribution!**
If start with the uniform distribution on a set $S$, then the probability of an event $E$ is

$$P(E) = \frac{|E|}{|S|}$$

When flipping $n$ fair coins what is the probability of getting exactly $k$ Hs?

A. $1/n$
B. $k/n$
C. $1/2^n$
D. $\binom{n}{k}/2^n$
E. None of the above.
Binomial distribution

When flipping $n$ fair coins what is the probability of getting exactly $k$ Hs?

\[ P(k \text{ Hs}) = \frac{\# \text{ coin toss sequences with } k \text{ Hs}}{\# \text{ possible coin toss sequences}} \]

Possible coin toss sequences: \{ HH..HH, HH..HT, ..., TT..TH, TT..TT \}

\[ P(k \text{ Hs}) = \frac{\binom{n}{k}}{2^n} \]
Binomial distribution

When flipping \( n \) fair coins what is the probability of getting exactly \( k \) Hs?

\[
P(k \text{ Hs}) = \frac{\# \text{ coin toss sequences with } k \text{ Hs}}{\# \text{ possible coin toss sequences}}
\]

Possible coin toss sequences: \{ HH..HH, HH..HT, ..., TT..TH, TT..TT \}

\[
P(k \text{ Hs}) = \frac{\binom{n}{k}}{2^n}
\]

What if the coin isn't fair?
**Bernoulli trial:** a performance of an experiment with two possible outcomes.  
* e.g. flipping a coin

**Binomial distribution:** probability of exactly k successes in n independent Bernoulli trials, when probability of success is p.  
* e.g. # Hs in n coin flips when probability of H is p

---

What is it?  
A. \(\frac{C(n,k)}{2^n}\)  
B. \(p^k/2^n\)  
C. \(C(n,k) p^k\)  
D. \(C(n,k) p^k (1-p)^{n-k}\)  
E. None of the above.

---

**Rosen p. 480**

\[
\begin{align*}
\text{k H’s in n coin flips, } & \quad \text{probs of } H \text{ each time } = p \quad \sqrt{\text{H’s}} \\
\frac{H}{\bar{H}} & \quad \frac{H}{\bar{H}} \\
\frac{1-p}{p} & \quad \frac{1-p}{p} \\
\frac{1-p}{p} & \quad \frac{1-p}{p} \\
(1-p) & \quad (1-p) \\
\binom{n}{k} & \quad (n) \\
\end{align*}
\]
Randomness in the world

For the Nate-haters, here’s the 538 prediction and actual results side by side pic.twitter.com/jbny4pRX

Michael Cosentino

15 hours ago

Nate Sliver: statistician famous for analyzing election predictions & baseball
Randomness in Computer Science

When the **input** is random …

- data mining
  - *elections*
  - *weather*
  - *stock prices*
  - *genetic markers*

- analyzing experimental data

*When is analysis valid? When are we overfitting to available data?*
Randomness in Computer Science

When the desired output is random …

- picking a cryptographic key
- performing a scientific simulation
- programming a computer adversary in a game
Randomness in Computer Science

When the **algorithm** uses randomness …

- Monte Carlo methods  *Rosen p. 463*

- search heuristics  *avoid local mins*

- randomized hashing

- quicksort
"Intuitive probabilistic reasoning" ....
The Monty Hall Puzzle

Car hidden behind one of three doors.

Goats hidden behind the other two.

Player gets to choose a door.

Host opens another door, reveals a goat.

Player can choose whether to swap choice with other closed door or stay with original choice.

What's the player's best strategy?
A. Always swap.  
B. Always stay.  
C. Doesn't matter, it's 50/50.
Some history…

Puzzle introduced by **Steve Selvin** in 1975.

**Marilyn vos Savant** was a prodigy with record scores on IQ tests who wrote an advice column. In 1990, a reader asked for the solution to the Monty Hall puzzle.

- After she published the (correct) answer, thousands of readers (including PhDs and even a professor of statistics) demanded that she correct her "mistake".
- She built a simulator to demonstrate the solution so they could see for themselves how it worked.
Pick a door at random to start

The Monty Hall Puzzle … the solution
The Monty Hall Puzzle … the solution

- Contestant chooses: Goat 1, Goat 2, Car
- Host shows: Goat 2, Goat 1, Goat 2, Goat 1

The probability of winning the car by switching doors is 2/3.
The Monty Hall Puzzle … the solution

What's the probability of winning (C) if always switch ("Y")?

A. 1/3
B. 1/2
C. 2/3
D. 1
E. None of the above.
The Monty Hall Puzzle … the solution

What's the probability of winning (C) if always stay ("N")?

A. 1/3  
B. 1/2  
C. 2/3  
D. 1  
E. None of the above.
What's wrong with the following argument?

"It doesn't matter whether you stay or swap because the host opened one door to show a goat so there are only two doors remaining, and both of them are equally likely to have the car because the prizes were placed behind the doors randomly at the start of the game"
Conditional probabilities

Probability of an event may **change** if have additional information about outcomes.

Suppose $E$ and $F$ are events, and $P(F)>0$. Then,

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

given that

i.e.

$$P(E \cap F) = P(E|F)P(F)$$

Rosen p. 456
Conditional probabilities

Are these probabilities equal?

The probability that **two siblings are girls** if know the oldest is a girl.

The probability that **two siblings are boys** if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

A. They're equal.
B. They're not equal.
C. ???
Are these probabilities equal?
The probability that **two siblings are girls** if know the oldest is a girl.
The probability that **two siblings are boys** if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

1. **Sample space**

2. **Initial distribution on the sample space**

3. **What events are we conditioning on?**
Are these probabilities equal?  
The probability that two siblings are girls if know the oldest is a girl.  
The probability that two siblings are boys if know that one of them is a boy.  

Assume that each child being a boy or a girl is equally likely.

1. Sample space  
   Possible outcomes: \{bb, bg, gb, gg\} \hspace{1cm} \text{Order matters!}

2. Initial distribution on the sample space

3. What events are we conditioning on?
Conditional probabilities

Are these probabilities equal?
The probability that two siblings are girls if know the oldest is a girl.
The probability that two siblings are boys if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

1. Sample space
   Possible outcomes: \{bb, bg, gb, gg\}  \textit{Order matters!}

2. Initial distribution on the sample space
   Uniform distribution, each outcome has probability \(\frac{1}{4}\).

3. What events are we conditioning on?
Are these probabilities equal?
The probability that **two siblings are girls** if know the oldest is a girl.
The probability that **two siblings are boys** if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

3. **What events are we conditioning on?**
Are these probabilities equal?
The probability that two siblings are girls if know the oldest is a girl.
The probability that two siblings are boys if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

3. What events are we conditioning on?
A = \{ outcomes where oldest is a girl \} \quad B = \{ outcomes where two are girls\}
Conditional probabilities

Are these probabilities equal?
The probability that two siblings are girls if know the oldest is a girl.
The probability that two siblings are boys if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

3. What events are we conditioning on?
A = { outcomes where oldest is a girl } = { gg, gb }
B = { outcomes where two are girls } = { gg }

\[ P(\text{boys} \mid A) \]
Conditional probabilities

Are these probabilities equal?
The probability that two siblings are girls if know the oldest is a girl.
The probability that two siblings are boys if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

3. What events are we conditioning on?
A = \{ outcomes where oldest is a girl \} = \{ gg, gb \}
P(A) = \frac{1}{2}

B = \{ outcomes where two are girls \} = \{ gg \}
P(B) = \frac{1}{4} = P(A \cap B)

\[ A \cap B = \{ gg \} \]
Are these probabilities equal?
The probability that two siblings are girls if know the oldest is a girl.
The probability that two siblings are boys if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

3. What events are we conditioning on?
A = \{ outcomes where oldest is a girl \} = \{ gg, gb \}
P(A) = \frac{1}{2}

B = \{ outcomes where two are girls \} = \{ gg \}
P(B) = \frac{1}{4} = P(A \cap B)

By conditional probability law: P(B | A) = P(A \cap B) / P(A) = (1/4) / (1/2) = \frac{1}{2}.
Conditional probabilities

Are these probabilities equal?
The probability that two siblings are girls if know the oldest is a girl. 1/2
The probability that two siblings are boys if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

1. Sample space
Possible outcomes: \{bb, bg, gb, gg\}  Order matters!

2. Initial distribution on the sample space
Uniform distribution, each outcome has probability 1/4.

3. What events are we conditioning on?
Conditional probabilities

Are these probabilities equal?
The probability that two siblings are girls if know the oldest is a girl. 1/2

The probability that two siblings are boys if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

3. What events are we conditioning on?
C = { outcomes where one is a boy}  D = { outcomes where two are boys }
   = { bb, bg, gb }                = { bb }
P(C) = 3/4                                      P(D) = 1/4 = P(C ∩ D)  

\[ \text{goal: } P(D|C) \]

\[ C \cap D = \{ bb \} \]
Conditional probabilities

Are these probabilities equal?
The probability that two siblings are girls if know the oldest is a girl. 1/2
The probability that two siblings are boys if know that one of them is a boy.

Assume that each child being a boy or a girl is equally likely.

3. What events are we conditioning on?
C = { outcomes where one is a boy} = { bb, bg, gb }
P(C) = 3/4

D = { outcomes where two are boys } = { bb }
P(D) = 1/4 = P(C ∩ D)

By conditional probability law: P(D | C) = P(C ∩ D) / P(C) = (1/4) / (3/4) = 1/3.
Conditional probabilities

Are these probabilities equal?
The probability that two siblings are girls if know the oldest is a girl. 1/2
The probability that two siblings are boys if know that one of them is a boy. 1/3

Assume that each child being a boy or a girl is equally likely.
### Conditional probabilities: Simpson's Paradox

<table>
<thead>
<tr>
<th></th>
<th>Treatment A</th>
<th>Treatment B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Small stones</strong></td>
<td>81 successes / 87 (93%)</td>
<td>234 successes / 270 (87%)</td>
</tr>
<tr>
<td><strong>Large stones</strong></td>
<td>192 successes / 263 (73%)</td>
<td>55 successes / 80 (69%)</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td>273 successes / 350 (78%)</td>
<td>289 successes / 350 (83%)</td>
</tr>
</tbody>
</table>

Which treatment is better?

A. Treatment A for all cases.  
B. Treatment B for all cases.  
C. A for small and B for large.  
D. A for large and B for small.

---

### Conditional probabilities: Simpson's Paradox

<table>
<thead>
<tr>
<th></th>
<th>Treatment A</th>
<th>Treatment B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Small stones</strong></td>
<td>81 successes / 87 (93%)</td>
<td>234 successes / 270 (87%)</td>
</tr>
<tr>
<td><strong>Large stones</strong></td>
<td>192 successes / 263 (73%)</td>
<td>55 successes / 80 (69%)</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td>273 successes / 350 (78%)</td>
<td>289 successes / 350 (83%)</td>
</tr>
</tbody>
</table>

**Simpson's Paradox**

"When the less effective treatment is applied more frequently to easier cases, it can appear to be a more effective treatment."

---

A random variable assigns a real number to each possible outcome of an experiment.

The distribution of a random variable $X$ is the function

$$ r \rightarrow P(X = r) $$

The expectation (average, expected value) of random variable $X$ on sample space $S$ is

$$ E(X) = \sum_{s \in S} P(s)X(s) $$

$$ = \frac{3}{8} \cdot 0 + \frac{3}{8} \cdot 1 + \frac{3}{8} \cdot 2 + \frac{1}{8} \cdot 3 $$

$$ = \frac{3}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} $$

$$ = \frac{3}{2} = 1.5 $$
The expectation (average, expected value) of random variable $X$ on sample space $S$ is

$$E(X) = \sum_{s \in S} P(s)X(s) = \sum_{r \in X(S)} P(X = r)r$$

More precisely: a fair coin is flipped twice. Let $S$ be the sample space of four possible outcomes. Let $X$ be the random variable that assigns to an outcome the number of heads in this outcome. What is $E(X)$?

What's the expected number of heads when we flip a coin twice?

A. 0

B. 1

C. 2

D. >2

E. None of the above.
The **expectation** (average, expected value) of random variable $X$ on sample space $S$ is

$$E(X) = \sum_{s \in S} P(s)X(s) = \sum_{r \in X(S)} P(X = r)r$$

More precisely: a fair coin is flipped three times. Let $S$ be the sample space of four possible outcomes. Let $X$ be the random variable that assigns to an outcome the number of heads in this outcome. What is $E(X)$?

What's the expected number of heads when we flip a coin three times?

A. 0
B. 1
C. 2
D. >2
E. None of the above.

$$= 1.5$$
The **expectation** (average, expected value) of random variable $X$ on sample space $S$ is

$$E(X) = \sum_{s \in S} P(s)X(s) = \sum_{r \in X(S)} P(X = r)r$$

- $E(X)$ may not be an actually possible value of $X$.
- $m \leq E(X) \leq M$, where $m$ is minimum value of $X$ and $M$ is maximum value of $X$. 

*Rosen p. 460,478*
Useful trick 1: Case analysis

The expectation can be computed by conditioning on an event and its complement

**Theorem:** For any random variable $X$ and event $A$,

$$E(X) = P(A) E(X | A) + P(A^c) E(X | A^c)$$

where $A^c$ is the complement of $A$. 

Rosen p. 460,478

Conditional Expectation
**Example:** If $X$ is the number of pairs of consecutive Hs when we flip a fair coin three times, what is the expectation of $X$?

e.g. $X(\text{HHT}) = 1$
$X(\text{HHH}) = 2$. 
Useful trick 1: Case analysis

**Example:** If $X$ is the number of pairs of consecutive Hs when we flip a fair coin three times, what is the expectation of $X$?

**Solution:**

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

*Directly from definition*

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

For each of eight possible outcomes, find probability and value of $X$:

HHH ($P(\text{HHH}) = 1/8$, $X(\text{HHH}) = 2$), HHT, HTH, HTT, THH, THT, TTH, TTT etc.
Useful trick 1: Case analysis

Example: If X is the number of pairs of consecutive Hs when we flip a fair coin three times, what is the expectation of X?

Solution:

Using conditional expectation

Let A be the event "The middle flip is H".

Which subset of S is A?
A. \{ HHH \}
B. \{ THT \}
C. \{ HHT, THH\}
D. \{ HHH, HHT, THH, THT\}
E. None of the above.
Useful trick 1: Case analysis

Example: If $X$ is the number of pairs of consecutive Hs when we flip a fair coin three times, what is the expectation of $X$?

Solution:

Using conditional expectation

Let $A$ be the event "The middle flip is H".

$$E(X) = P(A) E(X | A) + P( A^c ) E ( X | A^c )$$
Useful trick 1: Case analysis

Example: If $X$ is the number of pairs of consecutive Hs when we flip a fair coin three times, what is the expectation of $X$?

Solution: 

*Using conditional expectation*

Let $A$ be the event "The middle flip is H". 

$$P(A) = 1/2, \ P(A^c) = 1/2$$

$$E(X) = P(A) \ E(X \mid A) + P(A^c) \ E(X \mid A^c)$$
Useful trick 1: Case analysis

Example: If $X$ is the number of pairs of consecutive Hs when we flip a fair coin three times, what is the expectation of $X$?

Solution:

*Using conditional expectation*

Let $A$ be the event "The middle flip is H". $P(A) = 1/2$, $P(A^c) = 1/2$

$$E(X) = P(A)E(X | A) + P(A^c)E(X | A^c)$$

$E(X | A^c)$: If middle flip isn't H, there can't be any pairs of consecutive Hs
Example: If $X$ is the number of pairs of consecutive Hs when we flip a fair coin three times, what is the expectation of $X$?

Solution:

*Using conditional expectation*

Let $A$ be the event "The middle flip is H".  

$$P(A) = 1/2, \quad P(A^c) = 1/2$$

$$E(X) = P(A) E(X | A) + P(A^c) E(X | A^c)$$

$E(X | A^c)$: If middle flip isn't H, there can't be *any* pairs of consecutive Hs

$E(X | A)$: If middle flip is H, # pairs of consecutive Hs = # Hs in first & last flips
Example: If $X$ is the number of pairs of consecutive Hs when we flip a fair coin three times, what is the expectation of $X$?

Solution:

*Using conditional expectation*

Let $A$ be the event "The middle flip is H". 

$$E(X) = P(A)E(X | A) + P(A^c)E(X | A^c)$$

$E( X | A^c ) = 0$

$E( X | A ) = \frac{1}{4} \times 0 + \frac{1}{2} \times 1 + \frac{1}{4} \times 2 = 1$

$P(A) = 1/2 \ , \ P(A^c) = 1/2$
Example: If $X$ is the number of pairs of consecutive Hs when we flip a fair coin three times, what is the expectation of $X$?

Solution:

*Using conditional expectation*

Let $A$ be the event "The middle flip is H". 

\[
E(X) = P(A) E(X | A) + P(A^c) E(X | A^c) = \frac{1}{2} (1) + \frac{1}{2} (0) = 1/2
\]

\[
E(X | A^c) = 0
\]

\[
E(X | A) = \frac{1}{4} * 0 + \frac{1}{2} * 1 + \frac{1}{4} * 2 = 1
\]
Useful trick 1: Case analysis

Examples: Ending condition
- Each time I play solitaire I have a probability $p$ of winning. I play until I win a game.
- Each time a child is born, it has probability $p$ of being left-handed. I keep having kids until I have a left-handed one.

Let $X$ be the number of games OR number of kids until ending condition is met.

What's $E(X)$?

A. 1.
B. Some big number that depends on $p$.
C. $1/p$.
D. None of the above.
Useful trick 1: Case analysis

Ending condition

Let $X$ be the number of games OR number of kids until ending condition is met.

Solution:

*Directly from definition*

Need to compute the sum of all possible $P(X = i) i$. 

Useful trick 1: Case analysis

Ending condition

Let $X$ be the number of games OR number of kids until ending condition is met.

Solution:

Directly from definition

Need to compute the sum of all possible $P(X = i) i$.

$P(X = i) = \text{Probability that don't stop the first } i-1 \text{ times and do stop at the } i^{th} \text{ time}$

$= (1-p)^{i-1} p$
Useful trick 1: Case analysis

Ending condition

Let $X$ be the number of games OR number of kids until ending condition is met.

Solution:

*Directly from definition*

Need to compute the sum of all possible $P(X = i) i$.

$P(X = i) = \text{Probability that don't stop the first } i-1 \text{ times and do stop at the } i^{th} \text{ time} = (1-p)^{i-1} p$

$$E(x) = \sum_{i=1}^{\infty} i(1 - p)^{i-1} p$$
Useful trick 1: Case analysis

Ending condition

Let \( X \) be the number of games OR number of kids until ending condition is met.

Solution:

*Using conditional expectation*

\[
E(X) = P(A) \ E(X \mid A) + P(A^c) \ E(X \mid A^c)
\]
Useful trick 1: Case analysis

Ending condition

Let $X$ be the number of games OR number of kids until ending condition is met.

Solution:

*Using conditional expectation* Let $A$ be the event "success at first try".

$$E(X) = P(A) \cdot E(X | A) + P(A^c) \cdot E(X | A^c)$$
Useful trick 1: Case analysis

Ending condition

Let $X$ be the number of games OR number of kids until ending condition is met.

Solution:

Using conditional expectation

Let $A$ be the event "success at first try".

$$E(X) = P(A)E(X | A) + P(A^c)E(X | A^c)$$

$P(A) = p$  $P(A^c) = 1-p$
Useful trick 1: Case analysis

Ending condition

Let $X$ be the number of games OR number of kids until ending condition is met.

Solution:

Using conditional expectation

Let $A$ be the event "success at first try".

$$E(X) = P(A) E(X | A) + P(A^c) E(X | A^c)$$

$$P(A) = p \quad P(A^c) = 1-p$$

$$E(X|A) = 1 \quad because \ stop \ after \ first \ try$$
Useful trick 1: Case analysis

Ending condition

Let $X$ be the number of games OR number of kids until ending condition is met.

Solution:

*Using conditional expectation*

Let $A$ be the event "success at first try".

$$E(X) = P(A)\ E(X \mid A) + P(A^c)\ E(X \mid A^c)$$

$$P(A) = p \quad P(A^c) = 1-p$$

$$E(X \mid A) = 1$$

$$E(X \mid A^c) = 1 + E(X) \quad \text{because tried once and then at same situation from start}$$
Useful trick 1: Case analysis

Ending condition

Let $X$ be the number of games OR number of kids until ending condition is met.

Solution:

*Using conditional expectation*

Let $A$ be the event "success at first try".

$$E(X) = P(A) \cdot E(X \mid A) + P( A^c ) \cdot E ( X \mid A^c )$$

$P(A) = p$  \quad $P(A^c) = 1-p$

$E(X \mid A) = 1$

$E(X \mid A^c) = 1 + E(X)$

$$E(X) = p(1) + ( 1-p ) (1 + E(X) )$$
Useful trick 1: Case analysis

Ending condition

Let X be the number of games OR number of kids until ending condition is met.

Solution:

*Using conditional expectation* Let A be the event "success at first try".

\[
E(X) = p(1) + (1-p)(1 + E(X))
\]

Solving for \(E(X)\) gives:

\[
E(x) = \frac{1}{p}
\]
Useful trick 2: Linearity of expectation

**Theorem:** If $X_i$ are random variables on $S$ and if $a$ and $b$ are real numbers then

\[
E(X_1 + \ldots + X_n) = E(X_1) + \ldots + E(X_n)
\]

and

\[
E(aX + b) = aE(X) + b.
\]
Useful trick 2: Linearity of expectation

Example: Expected number of *consecutive heads* when we flip a fair coin $n$ times?

A. 1.
B. 2.
C. $n$.
D. None of the above.
E. ??

$$N = 3$$
Know answer $= \frac{1}{2}$
Useful trick 2: Linearity of expectation

Example: Expected number of consecutive heads when we flip a fair coin \( n \) times?

Solution: Define \( X_i = 1 \) if both the \( i^{th} \) and \( i+1^{st} \) flips are H; \( X_i=0 \) otherwise.

Looking for \( E(X) \) where

\[
X = \sum_{i=1}^{n-1} X_i.
\]

For each \( i \), what is \( E(X_i) \)?

A. \( 0 \).
B. \( \frac{1}{4} \).
C. \( \frac{1}{2} \).
D. 1.
E. ??

\[
\frac{1}{4}(1) + \frac{3}{4}(0)
\]
Example: Expected number of **consecutive heads** when we flip a fair coin \(n\) times?

Solution: Define \(X_i = 1\) if both the \(i^{th}\) and \((i+1)^{st}\) flips are H; \(X_i = 0\) otherwise.

Looking for \(E(X)\) where

\[
X = \sum_{i=1}^{n-1} X_i.
\]

\[
E(X) = \sum_{i=1}^{n-1} E(X_i) = \sum_{i=1}^{n-1} \frac{1}{4} = \frac{n - 1}{4}
\]

**Sanity check**

\[
\frac{3 - 1}{4} = \frac{2}{4} = \frac{1}{2}
\]

**Linearity - do each part separately**
Useful trick 2: Linearity of expectation

Example: Expected number of consecutive heads when we flip a fair coin $n$ times?

Solution: Define $X_i = 1$ if both the $i^{th}$ and $i+1^{st}$ flips are H; $X_i=0$ otherwise.

Looking for $E(X)$ where

$$X = \sum_{i=1}^{n-1} X_i.$$  

Indicator variables: 1 if pattern occurs, 0 otherwise

$$E(X) = \sum_{i=1}^{n-1} E(X_i) = \sum_{i=1}^{n-1} \frac{1}{4} = \frac{n - 1}{4}$$
Other functions?

Expectation does not in general commute with other functions.

\[ E( f(X) ) \neq f( E(X) ) \]

For example, let X be random variable with \( P(X = 0) = \frac{1}{2}, P(X = 1) = \frac{1}{2} \)

What's \( E(X) \)?

What's \( E(X^2) \)?

What's \( (E(X))^2 \)?
Other functions?

Expectation does not in general commute with other functions.

\[ E \left( f(X) \right) \neq f \left( E(X) \right) \]

For example, let \( X \) be random variable with \( P(X = 0) = \frac{1}{2} \), \( P(X = 1) = \frac{1}{2} \)

What's \( E(X) \)? \( \left( \frac{1}{2} \right)0 + \left( \frac{1}{2} \right)1 = \frac{1}{2} \)

What's \( E(X^2) \)? \( \left( \frac{1}{2} \right)0^2 + \left( \frac{1}{2} \right)1^2 = \frac{1}{2} \)

What's \( \left( E(X) \right)^2 \)? \( \left( \frac{1}{2} \right)^2 = \frac{1}{4} \)
Reminders

HW 8 due **Monday, Nov 30** at 11:59pm via Gradescope.

Save the date:
- Final exam is Saturday, Dec 5 at 11:30am.
- That’s the Saturday **before** Finals Week.