CS 13

Mathematical Foundations of Computing

Expectation

Random Variables

This is a variable:

```
variable = 17
```

This is a variable:

```
variable = 17
```

This is a **random variable**:

```
random_variable = RollDie(6)
```

Random Variables

This is a variable:

```
variable = 17
```

This is a **random variable**:

```
random_variable = RollDie(6)
```

These are also **random variable**s:

```
r_var_1 = RollDie(2)
r_var_2 = RollDie(3)
r_var_3 = r_var_1 + r_var_2
```

This is a variable:

```
variable = 17
```

This is a **random variable**:

```
random_variable = RollDie(6)
```

These are also **random variable**s:

```
r_var_1 = RollDie(2)
r_var_2 = RollDie(3)
r_var_3 = r_var_1 + r_var_2
```

Let Ω be the sample space of an experiment.

Random Variables

This is a variable:

```
variable = 17
```

This is a **random variable**:

```
random_variable = RollDie(6)
```

These are also **random variable**s:

```
r_var_1 = RollDie(2)
r_var_2 = RollDie(3)
r_var_3 = r_var_1 + r_var_2
```

Let Ω be the sample space of an experiment.
Formally, we can view a random variable X as a function from Ω to $\mathbb{N}$.

Random Variables

This is a variable:

```
variable = 17
```

This is a **random variable**:

```
random_variable = RollDie(6)
```

These are also **random variable**s:

```
r_var_1 = RollDie(2)
r_var_2 = RollDie(3)
r_var_3 = r_var_1 + r_var_2
```

Let Ω be the sample space of an experiment.
Formally, we can view a random variable X as a function from Ω to $\mathbb{N}$.
Looking at the examples above: $\texttt{r\_var\_1} : [2] \to [2]$

```
r_var_1 = RollDie(2)
r_var_2 = RollDie(3)
r_var_3 = r_var_1 + r_var_2
```

We often want to talk about the **probability mass function** of a random variable:

Random Variables


```
r_var_1 = RollDie(2)
r_var_2 = RollDie(3)
r_var_3 = r_var_1 + r_var_2
```

We often want to talk about the **probability mass function** of a random variable:

$$\Pr(\texttt{r\_var\_1} = x) = \frac{1}{2} \text{ (for } x \in [2]\text{)}$$

```
r_var_1 = RollDie(2)
r_var_2 = RollDie(3)
r_var_3 = r_var_1 + r_var_2
```

We often want to talk about the **probability mass function** of a random variable:

$$\Pr(\texttt{r\_var\_1} = x) = \frac{1}{2} \text{ (for } x \in [2])$$

$$\Pr(\texttt{r\_var\_2} = x) = \frac{1}{3} \text{ (for } x \in [3])$$

```
r_var_1 = RollDie(2)
r_var_2 = RollDie(3)
r_var_3 = r_var_1 + r_var_2
```

We often want to talk about the **probability mass function** of a random variable:

$$\Pr(\texttt{r\_var\_1} = x) = \frac{1}{2} \text{ (for } x \in [2])$$

$$\Pr(\texttt{r\_var\_2} = x) = \frac{1}{3} \text{ (for } x \in [3])$$

$$\Pr(\texttt{r\_var\_3} = x) = \begin{cases} 1/6 & \text{if } x = 2 \\ 2/6 & \text{if } x = 3 \\ 2/6 & \text{if } x = 4 \\ 1/6 & \text{if } x = 5 \end{cases}$$

```
r_var_1 = RollDie(2)
r_var_2 = RollDie(3)
r_var_3 = r_var_1 + r_var_2
```

We often want to talk about the **probability mass function** of a random variable:

$$\Pr(\texttt{r\_var\_1} = x) = \frac{1}{2} \text{ (for } x \in [2]\text{)}$$

$$\Pr(\texttt{r\_var\_2} = x) = \frac{1}{3} \text{ (for } x \in [3]\text{)}$$

$$\Pr(\texttt{r\_var\_3} = x) = \begin{cases} 1/6 & \text{if } x = 2 \\ 2/6 & \text{if } x = 3 \\ 2/6 & \text{if } x = 4 \\ 1/6 & \text{if } x = 5 \end{cases}$$

So, to recap, a random variable is a variable in a program that depends on a random process.

```
r_var_1 = RollDie(2)
r_var_2 = RollDie(3)
r_var_3 = r_var_1 + r_var_2
```

We often want to talk about the **probability mass function** of a random variable:

$$\Pr(\texttt{r\_var\_1} = x) = \frac{1}{2} \text{ (for } x \in [2])$$

$$\Pr(\texttt{r\_var\_2} = x) = \frac{1}{3} \text{ (for } x \in [3])$$

$$\Pr(\texttt{r\_var\_3} = x) = \begin{cases} 1/6 & \text{if } x = 2 \\ 2/6 & \text{if } x = 3 \\ 2/6 & \text{if } x = 4 \\ 1/6 & \text{if } x = 5 \end{cases}$$

So, to recap, a random variable is a variable in a program that depends on a random process. For our purposes, we will assume they always take on **natural number** values.

PVA 321 has the following grading policy:

- The final exam is worth 100% of the grade.

PVA 321 has the following grading policy:

- The final exam is worth 100% of the grade.

Suppose that we know the following facts:

Grading Policies: PVA 321

PVA 321 has the following grading policy:

- The final exam is worth 100% of the grade.

Suppose that we know the following facts:

- No student will ever get below 20.

PVA 321 has the following grading policy:

- The final exam is worth 100% of the grade.

Suppose that we know the following facts:

- No student will ever get below 20.
- Students are equally likely to get any score between 20 and 100.

Grading Policies: PVA 321

PVA 321 has the following grading policy:

- The final exam is worth 100% of the grade.

Suppose that we know the following facts:

- No student will ever get below 20.
- Students are equally likely to get any score between 20 and 100.

What is the average grade in PVA 321?

PVA 321 has the following grading policy:

- The final exam is worth 100% of the grade.

Suppose that we know the following facts:

- No student will ever get below 20.
- Students are equally likely to get any score between 20 and 100.

What is the average grade in PVA 321?

We know scores between 20 and 100 are equally likely; so, let's say they each show up k times.

PVA 321 has the following grading policy:

- The final exam is worth 100% of the grade.

Suppose that we know the following facts:

- No student will ever get below 20.
- Students are equally likely to get any score between 20 and 100.

What is the average grade in PVA 321?

We know scores between 20 and 100 are equally likely; so, let's say they each show up k times.

Now, just using the definition of average, we get

$$\frac{20k + 21k + \cdots + 100k}{81k} = \frac{1}{81}\left(\sum_{i=20}^{100} i\right)$$

Grading Policies: PVA 322

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

Grading Policies: PVA 322

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

- No student will ever get below 20.

Grading Policies: PVA 322

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

- No student will ever get below 20.
- 20% of the students got a 100.

Grading Policies: PVA 322

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Grading Policies: PVA 322

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class?

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class? Say there are n students total.

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class? Say there are n students total. Then, $0.2n$ of the students got a 100, $0.3n$ of the students got an 80, and for each score between 81 and 99, $\frac{0.5n}{19}$ students got that score.

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class? Say there are n students total. Then, $0.2n$ of the students got a 100, $0.3n$ of the students got an 80, and for each score between 81 and 99, $\frac{0.5n}{19}$ students got that score. So, taking the average gives us:

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class? Say there are n students total. Then, $0.2n$ of the students got a 100, $0.3n$ of the students got an 80, and for each score between 81 and 99, $\frac{0.5n}{19}$ students got that score. So, taking the average gives us:

$$\frac{100(0.2n)+80(0.3n)+\sum_{i=81}^{99} i\frac{0.5n}{19}}{n} = (100)(0.2)+(80)(0.3)+\sum_{i=81}^{99} i\frac{0.5}{19}$$

PVA 322 has the same grading policy:

- The final exam is worth 100% of the grade.

But the distribution of scores is different:

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class? Say there are n students total. Then, $0.2n$ of the students got a 100, $0.3n$ of the students got an 80, and for each score between 81 and 99, $\frac{0.5n}{19}$ students got that score. So, taking the average gives us:

$$\frac{100(0.2n) + 80(0.3n) + \sum_{i=81}^{99} i\frac{0.5n}{19}}{n} = (100)(0.2) + (80)(0.3) + \sum_{i=81}^{99} i\frac{0.5}{19}$$

What does this have to do with probability? Let's rephrase everything!

Grading Policies: PVA 322

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class? Let X be the r.v. for a student's score in PVA 322.

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class? Let X be the r.v. for a student's score in PVA 322. Then, the p.m.f. of X is:

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class? Let X be the r.v. for a student's score in PVA 322. Then, the p.m.f. of X is:

$$\Pr(X = i) = \begin{cases} 0.3 & \text{if } i = 80 \\ 0.5\frac{1}{19} & \text{if } 81 \leq i \leq 99 \\ 0.2 & \text{if } i = 100 \\ 0 & \text{otherwise} \end{cases}$$

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class? Let X be the r.v. for a student's score in PVA 322. Then, the p.m.f. of X is:

$$\Pr(X = i) = \begin{cases} 0.3 & \text{if } i = 80 \\ 0.5\frac{1}{19} & \text{if } 81 \le i \le 99 \\ 0.2 & \text{if } i = 100 \\ 0 & \text{otherwise} \end{cases}$$

To get the average, we just multiply each outcome by how likely it is:

- No student will ever get below 20.
- 20% of the students got a 100.
- 30% of the students got an 80.
- The remaining 50% of the students got scores evenly distributed between $81, 82, \ldots, 99$.

Same Question: What is the average score in the class? Let X be the r.v. for a student's score in PVA 322. Then, the p.m.f. of X is:

$$\Pr(X = i) = \begin{cases} 0.3 & \text{if } i = 80 \\ 0.5\frac{1}{19} & \text{if } 81 \le i \le 99 \\ 0.2 & \text{if } i = 100 \\ 0 & \text{otherwise} \end{cases}$$

To get the average, we just multiply each outcome by how likely it is:

$$\sum_{i=0}^{100} i\Pr(X = i) = 100\Pr(X = 100) + 80\Pr(X = 80) + \sum_{i=81}^{99} i\Pr(X = i)$$

$$= (100)(0.2) + (80)(0.3) + \sum_{i=81}^{99} i\frac{0.5}{19}$$

Definition (Random Variable)

A Little More Formal Now...

Definition (Random Variable)

A **random variable** is a variable in a randomized piece of code **or**
A **random variable** is a function from the sample space to the naturals.

Definition (Random Variable)

A **random variable** is a variable in a randomized piece of code **or**
A **random variable** is a function from the sample space to the naturals.

Definition (Expected Value)

The **expected value** of a random variable $X : \Omega \to \mathbb{N}$ is the average value it takes on.

Definition (Random Variable)

A **random variable** is a variable in a randomized piece of code **or**
A **random variable** is a function from the sample space to the naturals.

Definition (Expected Value)

The **expected value** of a random variable $X : \Omega \to \mathbb{N}$ is the average value it takes on. As we figured out above, we can calculate it with:

A Little More Formal Now...

Definition (Random Variable)

A **random variable** is a variable in a randomized piece of code **or**
A **random variable** is a function from the sample space to the naturals.

Definition (Expected Value)

The **expected value** of a random variable $X : \Omega \to \mathbb{N}$ is the average value it takes on. As we figured out above, we can calculate it with:

$$\mathbb{E}[X] = \sum_{x \in \Omega} X(x) \Pr(X = x) = \sum_{n=0}^{\infty} n \Pr(X = n)$$

Let X be the r.v. for the number of heads when we flip 2 coins with bias p. What is $\mathbb{E}[X]$?

Let X be the r.v. for the number of heads when we flip 2 coins with bias p. What is $\mathbb{E}[X]$?

Proof.

By definition of expected value, we know $\mathbb{E}[X] = \sum_{k=0}^{\infty} k \Pr(X = k)$.

Let X be the r.v. for the number of heads when we flip 2 coins with bias p. What is $\mathbb{E}[X]$?

Proof.

By definition of expected value, we know $\mathbb{E}[X] = \sum_{k=0}^{\infty} k \mathrm{Pr}(X = k)$.

We know that if $0 \le k \le n$, then

$$\mathrm{Pr}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Let X be the r.v. for the number of heads when we flip 2 coins with bias p. What is $\mathbb{E}[X]$?

Proof.

By definition of expected value, we know $\mathbb{E}[X] = \sum_{k=0}^{\infty} k\Pr(X = k)$.

We know that if $0 \leq k \leq n$, then

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

So,

$$\begin{aligned}\mathbb{E}[X] &= 0\binom{2}{0}p^0(1-p)^{2-0} + 1\binom{2}{1}p^1(1-p)^{2-1} + 2\binom{2}{2}p^2(1-p)^{2-2} \\ &= 2p(1-p) + 2p^2 \\ &= 2p\end{aligned}$$

□

Let X be the r.v. for the number of heads when we flip 2 coins with bias p. What is $\mathbb{E}[X]$?

Proof.

By definition of expected value, we know $\mathbb{E}[X] = \sum_{k=0}^{\infty} k\Pr(X = k)$.

We know that if $0 \le k \le n$, then

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

So,

$$\mathbb{E}[X] = 0\binom{2}{0}p^0(1-p)^{2-0} + 1\binom{2}{1}p^1(1-p)^{2-1} + 2\binom{2}{2}p^2(1-p)^{2-2}$$
$$= 2p(1-p) + 2p^2$$
$$= 2p$$

□

Let X be the r.v. for the number of heads when we flip 2 coins with bias p. What is $\mathbb{E}[X]$?

Proof.

By definition of expected value, we know $\mathbb{E}[X] = \sum_{k=0}^{\infty} k\Pr(X = k)$.

We know that if $0 \leq k \leq n$, then

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

So,

$$\begin{aligned}\mathbb{E}[X] &= 0\binom{2}{0}p^0(1-p)^{2-0} + 1\binom{2}{1}p^1(1-p)^{2-1} + 2\binom{2}{2}p^2(1-p)^{2-2} \\ &= 2p(1-p) + 2p^2 \\ &= 2p\end{aligned}$$

□

Let X be the r.v. for the number of heads when we flip 2 coins with bias p. What is $\mathbb{E}[X]$?

Proof.

By definition of expected value, we know $\mathbb{E}[X] = \sum_{k=0}^{\infty} k\Pr(X = k)$.

We know that if $0 \le k \le n$, then

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

So,

$$\begin{aligned}\mathbb{E}[X] &= 0\binom{2}{0}p^0(1-p)^{2-0} + 1\binom{2}{1}p^1(1-p)^{2-1} + 2\binom{2}{2}p^2(1-p)^{2-2} \\ &= 2p(1-p) + 2p^2 \\ &= 2p\end{aligned}$$

□

Check Up!

Consider the following expectations:

- $\mathbb{E}[1313]$
- $\mathbb{E}[$getting `HEADS` on a coin flip$]$
- $\mathbb{E}[$number of `HEADS` from a coin flip given that the flip is `HEADS`$]$

What are they?

Check Up!

Consider the following expectations:

- $\mathbb{E}[1313]$
- $\mathbb{E}[\text{getting } \texttt{HEADS} \text{ on a coin flip}]$
- $\mathbb{E}[\text{number of } \texttt{HEADS} \text{ from a coin flip given that the flip is } \texttt{HEADS}]$

What are they?

- If 1313 is a r.v., then it must be the one that is always 1313. So, $\mathbb{E}[1313] = 1313$.

Consider the following expectations:

- $\mathbb{E}[1313]$
- $\mathbb{E}[$getting `HEADS` on a coin flip$]$
- $\mathbb{E}[$number of `HEADS` from a coin flip given that the flip is `HEADS`$]$

What are they?

- If 1313 is a r.v., then it must be the one that is always 1313. So, $\mathbb{E}[1313] = 1313$.
- THIS DOESN'T MAKE SENSE! We can only take the expectation of a r.v.–not an event!

Check Up!

Consider the following expectations:

- $\mathbb{E}[1313]$
- $\mathbb{E}$[getting HEADS on a coin flip]
- $\mathbb{E}$[number of HEADS from a coin flip given that the flip is HEADS]

What are they?

- If 1313 is a r.v., then it must be the one that is always 1313. So, $\mathbb{E}[1313] = 1313$.
- THIS DOESN'T MAKE SENSE! We can only take the expectation of a r.v.–not an event!
- It's 1, because we're told the coin flip is HEADS

PVA 323 has the following grading policy:

- The midterm is worth 20% of the grade.
- The homework is worth 30% of the grade.
- The final exam is worth 50% of the grade.

PVA 323 has the following grading policy:

- The midterm is worth 20% of the grade.
- The homework is worth 30% of the grade.
- The final exam is worth 50% of the grade.

Stupid Question: Supposing that m was the average on the midterm, h was the average on the homework, and f was the average on the final exam, how do you calculate the average for the course?

PVA 323 has the following grading policy:

- The midterm is worth 20% of the grade.
- The homework is worth 30% of the grade.
- The final exam is worth 50% of the grade.

Stupid Question: Supposing that m was the average on the midterm, h was the average on the homework, and f was the average on the final exam, how do you calculate the average for the course?
Answer: $0.2m + 0.3h + 0.5f$.

PVA 323 has the following grading policy:

- The midterm is worth 20% of the grade.
- The homework is worth 30% of the grade.
- The final exam is worth 50% of the grade.

Stupid Question: Supposing that m was the average on the midterm, h was the average on the homework, and f was the average on the final exam, how do you calculate the average for the course?
Answer: $0.2m + 0.3h + 0.5f$.

Remember our interpretation of these exams as random variables?

More Grading Policies

PVA 323 has the following grading policy:

- The midterm is worth 20% of the grade.
- The homework is worth 30% of the grade.
- The final exam is worth 50% of the grade.

Stupid Question: Supposing that m was the average on the midterm, h was the average on the homework, and f was the average on the final exam, how do you calculate the average for the course?
Answer: $0.2m + 0.3h + 0.5f$.

Remember our interpretation of these exams as random variables? What we're actually saying here is if G is the r.v. for a PVA 323 student's course grade,

PVA 323 has the following grading policy:

- The midterm is worth 20% of the grade.
- The homework is worth 30% of the grade.
- The final exam is worth 50% of the grade.

Stupid Question: Supposing that m was the average on the midterm, h was the average on the homework, and f was the average on the final exam, how do you calculate the average for the course?
Answer: $0.2m+0.3h+0.5f$.

Remember our interpretation of these exams as random variables? What we're actually saying here is if G is the r.v. for a PVA 323 student's course grade, and M, H, and F are the obvious r.v.'s, then:

PVA 323 has the following grading policy:

- The midterm is worth 20% of the grade.
- The homework is worth 30% of the grade.
- The final exam is worth 50% of the grade.

Stupid Question: Supposing that m was the average on the midterm, h was the average on the homework, and f was the average on the final exam, how do you calculate the average for the course?
Answer: $0.2m + 0.3h + 0.5f$.

Remember our interpretation of these exams as random variables? What we're actually saying here is if G is the r.v. for a PVA 323 student's course grade, and M, H, and F are the obvious r.v.'s, then:

We know that $G = 0.2M + 0.3H + 0.5F$; so, we can conclude

PVA 323 has the following grading policy:

- The midterm is worth 20% of the grade.
- The homework is worth 30% of the grade.
- The final exam is worth 50% of the grade.

Stupid Question: Supposing that m was the average on the midterm, h was the average on the homework, and f was the average on the final exam, how do you calculate the average for the course?
Answer: $0.2m + 0.3h + 0.5f$.

Remember our interpretation of these exams as random variables? What we're actually saying here is if G is the r.v. for a PVA 323 student's course grade, and M, H, and F are the obvious r.v.'s, then:

We know that $G = 0.2M + 0.3H + 0.5F$; so, we can conclude

$$\mathbb{E}[G] = 0.2\mathbb{E}[M] + 0.3\mathbb{E}[H] + 0.5\mathbb{E}[F]$$

PVA 323 has the following grading policy:

- The midterm is worth 20% of the grade.
- The homework is worth 30% of the grade.
- The final exam is worth 50% of the grade.

Stupid Question: Supposing that m was the average on the midterm, h was the average on the homework, and f was the average on the final exam, how do you calculate the average for the course?
Answer: $0.2m + 0.3h + 0.5f$.

Remember our interpretation of these exams as random variables? What we're actually saying here is if G is the r.v. for a PVA 323 student's course grade, and M, H, and F are the obvious r.v.'s, then:

We know that $G = 0.2M + 0.3H + 0.5F$; so, we can conclude

$$\mathbb{E}[G] = 0.2\mathbb{E}[M] + 0.3\mathbb{E}[H] + 0.5\mathbb{E}[F]$$

This is incredibly powerful!

PVA 323 has the following grading policy:

- The midterm is worth 20% of the grade.
- The homework is worth 30% of the grade.
- The final exam is worth 50% of the grade.

Stupid Question: Supposing that m was the average on the midterm, h was the average on the homework, and f was the average on the final exam, how do you calculate the average for the course?
Answer: $0.2m + 0.3h + 0.5f$.

Remember our interpretation of these exams as random variables? What we're actually saying here is if G is the r.v. for a PVA 323 student's course grade, and M, H, and F are the obvious r.v.'s, then:

We know that $G = 0.2M + 0.3H + 0.5F$; so, we can conclude

$$\mathbb{E}[G] = 0.2\mathbb{E}[M] + 0.3\mathbb{E}[H] + 0.5\mathbb{E}[F]$$

This is incredibly powerful! To calculate the course average, we just added together other averages.

Definition (Linearity of Expectation)

If X, Y, and Z are r.v.'s such that $Z = X + Y$, then:

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Linearity of Expectation

Definition (Linearity of Expectation)

If X, Y, and Z are r.v.'s such that $Z = X + Y$, then:

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Generalizing a little, if we have $X = \sum_{i=1}^{n} X_i$, then

Definition (Linearity of Expectation)

If X, Y, and Z are r.v.'s such that $Z = X + Y$, then:

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Generalizing a little, if we have $X = \sum_{i=1}^{n} X_i$, then

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Definition (Linearity of Expectation)

If X, Y, and Z are r.v.'s such that $Z = X + Y$, then:

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Generalizing a little, if we have $X = \sum_{i=1}^{n} X_i$, then

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Definition (Indicator Random Variable)

Definition (Linearity of Expectation)

If X, Y, and Z are r.v.'s such that $Z = X + Y$, then:

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Generalizing a little, if we have $X = \sum_{i=1}^{n} X_i$, then

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Definition (Indicator Random Variable)

An **indicator random variable** "indicates" whether a particular event E happens or not.

Definition (Linearity of Expectation)

If X, Y, and Z are r.v.'s such that $Z = X + Y$, then:

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Generalizing a little, if we have $X = \sum_{i=1}^{n} X_i$, then

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Definition (Indicator Random Variable)

An **indicator random variable** "indicates" whether a particular event E happens or not. Usually, we define them via:

$$X = \begin{cases} 1 & \text{if BLAH happens} \\ 0 & \text{otherwise} \end{cases}$$

Definition (Linearity of Expectation)

If X, Y, and Z are r.v.'s such that $Z = X + Y$, then:

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Generalizing a little, if we have $X = \sum_{i=1}^{n} X_i$, then

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Definition (Indicator Random Variable)

An **indicator random variable** "indicates" whether a particular event E happens or not. Usually, we define them via:

$$X = \begin{cases} 1 & \text{if BLAH happens} \\ 0 & \text{otherwise} \end{cases}$$

They come up with Linearity of Expectation **a lot**!

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back.

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i \Pr(X = i)$.

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i \Pr(X = i)$. So, now we calculate $\Pr(X = i)$.

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i \Pr(X = i)$. So, now we calculate $\Pr(X = i)$.

$\Pr(X = i) =$

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i\Pr(X = i)$. So, now we calculate $\Pr(X = i)$.

$\Pr(X = i) = \binom{n}{i}$

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i\Pr(X = i)$. So, now we calculate $\Pr(X = i)$.

$\Pr(X = i) = \binom{n}{i}\frac{1}{n}$

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i \Pr(X = i)$. So, now we calculate $\Pr(X = i)$.

$\Pr(X = i) = \binom{n}{i} \frac{1}{n} \frac{1}{n-1} \cdots \frac{1}{n-i}$

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i \Pr(X = i)$. So, now we calculate $\Pr(X = i)$.

$\Pr(X = i) = \binom{n}{i}\frac{1}{n}\frac{1}{n-1}\cdots\frac{1}{n-i}\Pr(\text{none of the others get their hats back}))$.

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i \Pr(X = i)$. So, now we calculate $\Pr(X = i)$.

$\Pr(X = i) = \binom{n}{i} \frac{1}{n} \frac{1}{n-1} \cdots \frac{1}{n-i} \Pr(\text{none of the others get their hats back}))$.
I don't want to do this any more… make it easier?

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i \Pr(X = i)$. So, now we calculate $\Pr(X = i)$.
$\Pr(X = i) = \binom{n}{i} \frac{1}{n} \frac{1}{n-1} \cdots \frac{1}{n-i} \Pr(\text{none of the others get their hats back}))$.
I don't want to do this any more... make it easier?

Let X be the r.v. for the number of people who get their hat back.

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i\Pr(X = i)$. So, now we calculate $\Pr(X = i)$.
$\Pr(X = i) = \binom{n}{i}\frac{1}{n}\frac{1}{n-1}\cdots\frac{1}{n-i}\Pr(\text{none of the others get their hats back}))$.
I don't want to do this any more... make it easier?

Let X be the r.v. for the number of people who get their hat back.
Let X_i be the i.r.v. for the ith person getting his hat back.

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i\Pr(X = i)$. So, now we calculate $\Pr(X = i)$.
$\Pr(X = i) = \binom{n}{i}\frac{1}{n}\frac{1}{n-1}\cdots\frac{1}{n-i}\Pr(\text{none of the others get their hats back}))$.
I don't want to do this any more...make it easier?

Let X be the r.v. for the number of people who get their hat back.
Let X_i be the i.r.v. for the ith person getting his hat back.
Note that $X = \sum_{i=1}^{n} X_i$, because each of the n people either get their hat back or not.

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i\Pr(X = i)$. So, now we calculate $\Pr(X = i)$.
$\Pr(X = i) = \binom{n}{i}\frac{1}{n}\frac{1}{n-1}\cdots\frac{1}{n-i}\Pr(\text{none of the others get their hats back}))$.
I don't want to do this any more... make it easier?

Let X be the r.v. for the number of people who get their hat back.
Let X_i be the i.r.v. for the ith person getting his hat back.
Note that $X = \sum_{i=1}^{n} X_i$, because each of the n people either get their hat back or not. Now, consider
$\mathbb{E}[X_i] = 0\Pr(X_i = 0) + 1\Pr(X_i = 1) = \Pr(X_i = 1) = \frac{1}{n}$.

There are n people at a party who have checked their hats in. At the end of the party, the hat-check clerk randomly gives them their hats back. What is the expected number of people who got their hat back?

Let X be the r.v. for the number of people who get their hat back. By definition, $\mathbb{E}[X] = \sum_{i=0}^{n} i \Pr(X = i)$. So, now we calculate $\Pr(X = i)$.
$\Pr(X = i) = \binom{n}{i} \frac{1}{n} \frac{1}{n-1} \cdots \frac{1}{n-i} \Pr(\text{none of the others get their hats back}))$.
I don't want to do this any more... make it easier?

Let X be the r.v. for the number of people who get their hat back.
Let X_i be the i.r.v. for the ith person getting his hat back.
Note that $X = \sum_{i=1}^{n} X_i$, because each of the n people either get their hat back or not. Now, consider
$\mathbb{E}[X_i] = 0\Pr(X_i = 0) + 1\Pr(X_i = 1) = \Pr(X_i = 1) = \frac{1}{n}$.
So, $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} 1/n = 1$ by Linearity of Expectation.

Lots of Coins Again

What is the expected number of HEADS from flipping a coin (with bias p) n times?

What is the expected number of HEADS from flipping a coin (with bias p) n times? Let X be r.v. for the number of HEADS from flipping a coin (with bias p) n times.

What is the expected number of HEADS from flipping a coin (with bias p) n times? Let X be r.v. for the number of HEADS from flipping a coin (with bias p) n times.

We know $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

What is the expected number of HEADS from flipping a coin (with bias p) n times? Let X be r.v. for the number of HEADS from flipping a coin (with bias p) n times.

We know $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Putting this into our formula gives:
$\mathbb{E}[X] = \sum_{i=0}^{n} i \binom{n}{k} p^i (1-p)^{n-i}$.

Lots of Coins Again

What is the expected number of HEADS from flipping a coin (with bias p) n times? Let X be r.v. for the number of HEADS from flipping a coin (with bias p) n times.

We know $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Putting this into our formula gives: $\mathbb{E}[X] = \sum_{i=0}^{n} i \binom{n}{k} p^i (1-p)^{n-i}$. Any ideas on how to simplify this mess?

What is the expected number of HEADS from flipping a coin (with bias p) n times? Let X be r.v. for the number of HEADS from flipping a coin (with bias p) n times.

We know $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Putting this into our formula gives: $\mathbb{E}[X] = \sum_{i=0}^{n} i \binom{n}{k} p^i (1-p)^{n-i}$. Any ideas on how to simplify this mess?

Linearity of Expectation to the rescue!

What is the expected number of HEADS from flipping a coin (with bias p) n times? Let X be r.v. for the number of HEADS from flipping a coin (with bias p) n times.

We know $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Putting this into our formula gives:
$\mathbb{E}[X] = \sum_{i=0}^{n} i \binom{n}{k} p^i (1-p)^{n-i}$. Any ideas on how to simplify this mess?

Linearity of Expectation to the rescue!

Let X_i be the i.r.v. for the ith coin coming up HEADS.

What is the expected number of HEADS from flipping a coin (with bias p) n times? Let X be r.v. for the number of HEADS from flipping a coin (with bias p) n times.

We know $\Pr(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$. Putting this into our formula gives: $\mathbb{E}[X] = \sum_{i=0}^{n} i \binom{n}{k} p^i (1-p)^{n-i}$. Any ideas on how to simplify this mess?

Linearity of Expectation to the rescue!

Let X_i be the i.r.v. for the ith coin coming up HEADS. We know $\mathbb{E}[X_i] = \Pr(X_i = 1) = p$.

What is the expected number of HEADS from flipping a coin (with bias p) n times? Let X be r.v. for the number of HEADS from flipping a coin (with bias p) n times.

We know $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Putting this into our formula gives: $\mathbb{E}[X] = \sum_{i=0}^{n} i \binom{n}{k} p^i (1-p)^{n-i}$. Any ideas on how to simplify this mess?

Linearity of Expectation to the rescue!

Let X_i be the i.r.v. for the ith coin coming up HEADS. We know $\mathbb{E}[X_i] = \Pr(X_i = 1) = p$. So, since $X = \sum_{i=1}^{n} X_i$, we know,

What is the expected number of HEADS from flipping a coin (with bias p) n times? Let X be r.v. for the number of HEADS from flipping a coin (with bias p) n times.

We know $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Putting this into our formula gives: $\mathbb{E}[X] = \sum_{i=0}^{n} i \binom{n}{k} p^i (1-p)^{n-i}$. Any ideas on how to simplify this mess?

Linearity of Expectation to the rescue!

Let X_i be the i.r.v. for the ith coin coming up HEADS. We know $\mathbb{E}[X_i] = \Pr(X_i = 1) = p$. So, since $X = \sum_{i=1}^{n} X_i$, we know, by Linearity of Expectation, that $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$

What is the expected number of HEADS from flipping a coin (with bias p) n times? Let X be r.v. for the number of HEADS from flipping a coin (with bias p) n times.

We know $\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Putting this into our formula gives: $\mathbb{E}[X] = \sum_{i=0}^{n} i \binom{n}{k} p^i (1-p)^{n-i}$. Any ideas on how to simplify this mess?

Linearity of Expectation to the rescue!

Let X_i be the i.r.v. for the ith coin coming up HEADS. We know $\mathbb{E}[X_i] = \Pr(X_i = 1) = p$. So, since $X = \sum_{i=1}^{n} X_i$, we know, by Linearity of Expectation, that $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} p = np$.

Remember from last time, we mentioned that if $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$$

Law of Total Probability

Remember from last time, we mentioned that if $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$$

Let A and B be events.

Remember from last time, we mentioned that if $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$$

Let A and B be events. Note that $A \cap B$ and $A \cap \overline{B}$ are disjoint.

Remember from last time, we mentioned that if $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$$

Let A and B be events. Note that $A \cap B$ and $A \cap \overline{B}$ are disjoint. So,
$\Pr(A) = \Pr\big((A \cap B) \cup (A \cap \overline{B})\big)$

Law of Total Probability

Remember from last time, we mentioned that if $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$$

Let A and B be events. Note that $A \cap B$ and $A \cap \overline{B}$ are disjoint. So,
$\Pr(A) = \Pr\big((A \cap B) \cup (A \cap \overline{B})\big) = \Pr(A \cap B) + \Pr\big(A \cap \overline{B}\big)$.

Law of Total Probability

Remember from last time, we mentioned that if $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$$

Let A and B be events. Note that $A \cap B$ and $A \cap \overline{B}$ are disjoint. So,
$\Pr(A) = \Pr\big((A \cap B) \cup (A \cap \overline{B})\big) = \Pr(A \cap B) + \Pr\big(A \cap \overline{B}\big)$.
Using the definition of conditional probability, it follows that:

Law of Total Probability

Remember from last time, we mentioned that if $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$$

Let A and B be events. Note that $A \cap B$ and $A \cap \overline{B}$ are disjoint. So, $\Pr(A) = \Pr\big((A \cap B) \cup (A \cap \overline{B})\big) = \Pr(A \cap B) + \Pr\big(A \cap \overline{B}\big)$.
Using the definition of conditional probability, it follows that:

Definition (Law of Total Probability)

$$\Pr(A) = \Pr(A \mid B)\Pr(B) + \Pr\big(A \mid \overline{B}\big)\Pr\big(\overline{B}\big)$$

This can be very useful when calculating probabilities.

Remember from last time, we mentioned that if $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$$

Let A and B be events. Note that $A \cap B$ and $A \cap \overline{B}$ are disjoint. So, $\Pr(A) = \Pr\big((A \cap B) \cup (A \cap \overline{B})\big) = \Pr(A \cap B) + \Pr\big(A \cap \overline{B}\big)$.
Using the definition of conditional probability, it follows that:

Definition (Law of Total Probability)

$$\Pr(A) = \Pr(A \mid B)\Pr(B) + \Pr\big(A \mid \overline{B}\big)\Pr\big(\overline{B}\big)$$

This can be very useful when calculating probabilities. It has an analog for expectation:

Law of Total Probability

Remember from last time, we mentioned that if $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$$

Let A and B be events. Note that $A \cap B$ and $A \cap \overline{B}$ are disjoint. So,
$\Pr(A) = \Pr\big((A \cap B) \cup (A \cap \overline{B})\big) = \Pr(A \cap B) + \Pr\big(A \cap \overline{B}\big)$.
Using the definition of conditional probability, it follows that:

Definition (Law of Total Probability)

$$\Pr(A) = \Pr(A \mid B)\Pr(B) + \Pr\big(A \mid \overline{B}\big)\Pr\big(\overline{B}\big)$$

This can be very useful when calculating probabilities. It has an analog for expectation: Let X be a r.v.

Remember from last time, we mentioned that if $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$$

Let A and B be events. Note that $A \cap B$ and $A \cap \overline{B}$ are disjoint. So,
$\Pr(A) = \Pr\big((A \cap B) \cup (A \cap \overline{B})\big) = \Pr(A \cap B) + \Pr\big(A \cap \overline{B}\big)$.
Using the definition of conditional probability, it follows that:

Definition (Law of Total Probability)

$$\Pr(A) = \Pr(A \mid B)\Pr(B) + \Pr\big(A \mid \overline{B}\big)\Pr\big(\overline{B}\big)$$

This can be very useful when calculating probabilities. It has an analog for expectation: Let X be a r.v.

Definition (Law of Total Expectation)

$$\mathbb{E}[X] = \mathbb{E}[X \mid A]\Pr(A) + \mathbb{E}\big[X \mid \overline{A}\big]\Pr\big(\overline{A}\big)$$

Back to Last Lecture


```
X = 1
while FlipCoin(p) != HEADS:
    X++
```

Let X be the r.v. for how many loop iterations the code gets through.

```
X = 1
while FlipCoin(p) != HEADS:
   X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

```
X = 1
while FlipCoin(p) != HEADS:
  X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$\mathbb{E}[X] = \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}[X \mid \overline{H}]\Pr(\overline{H}) \quad \text{[Law of Total Expectation]}$$
$$= \mathbb{E}[X \mid H]\,p + \mathbb{E}[X \mid \overline{H}](1-p) \quad \text{[The coin has bias } p\text{]}$$
$$= p + \mathbb{E}[X \mid \overline{H}](1-p) \quad \text{[If we get HEADS, we're done!]}$$
$$= p + \mathbb{E}[X+1](1-p) \quad [\star]$$
$$= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) \quad \text{[Linearity of Expectation]}$$
$$= 1 + (1-p)\mathbb{E}[X] \quad \text{[Simplifying]}$$
$$= \frac{1}{p} \quad \text{[Solve for } \mathbb{E}[X]\text{]}$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}[X \mid \overline{H}] = \mathbb{E}[X+1]$.

Back to Last Lecture


```
X = 1
while FlipCoin(p) != HEADS:
    X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$
\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) && \text{[Law of Total Expectation]} \\
&= \mathbb{E}[X \mid H]\,p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[The coin has bias } p\text{]} \\
&= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[If we get HEADS, we're done!]} \\
&= p + \mathbb{E}[X+1](1-p) && [\star] \\
&= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) && \text{[Linearity of Expectation]} \\
&= 1 + (1-p)\mathbb{E}[X] && \text{[Simplifying]} \\
&= \frac{1}{p} && \text{[Solve for } \mathbb{E}[X]\text{]}
\end{aligned}
$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
   X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) && \text{[Law of Total Expectation]} \\
&= \mathbb{E}[X \mid H]\,p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[The coin has bias } p\text{]} \\
&= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[If we get HEADS, we're done!]} \\
&= p + \mathbb{E}[X+1](1-p) && [\star] \\
&= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) && \text{[Linearity of Expectation]} \\
&= 1 + (1-p)\mathbb{E}[X] && \text{[Simplifying]} \\
&= \frac{1}{p} && \text{[Solve for } \mathbb{E}[X]\text{]}
\end{aligned}$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

Back to Last Lecture


```
X = 1
while FlipCoin(p) != HEADS:
   X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$
\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) && \text{[Law of Total Expectation]} \\
&= \mathbb{E}[X \mid H]\,p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[The coin has bias } p\text{]} \\
&= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[If we get HEADS, we're done!]} \\
&= p + \mathbb{E}[X+1](1-p) && [\star] \\
&= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) && \text{[Linearity of Expectation]} \\
&= 1 + (1-p)\mathbb{E}[X] && \text{[Simplifying]} \\
&= \frac{1}{p} && \text{[Solve for } \mathbb{E}[X]\text{]}
\end{aligned}
$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
   X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$
\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) && \text{[Law of Total Expectation]} \\
&= \mathbb{E}[X \mid H]\, p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[The coin has bias } p\text{]} \\
&= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[If we get HEADS, we're done!]} \\
&= p + \mathbb{E}[X+1](1-p) && [\star] \\
&= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) && \text{[Linearity of Expectation]} \\
&= 1 + (1-p)\mathbb{E}[X] && \text{[Simplifying]} \\
&= \frac{1}{p} && \text{[Solve for } \mathbb{E}[X]\text{]}
\end{aligned}
$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
    X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$
\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X \mid H]\,\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) && \text{[Law of Total Expectation]} \\
&= \mathbb{E}[X \mid H]\,p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[The coin has bias } p\text{]} \\
&= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[If we get HEADS, we're done!]} \\
&= p + \mathbb{E}[X+1](1-p) && [\star] \\
&= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) && \text{[Linearity of Expectation]} \\
&= 1 + (1-p)\mathbb{E}[X] && \text{[Simplifying]} \\
&= \frac{1}{p} && \text{[Solve for } \mathbb{E}[X]\text{]}
\end{aligned}
$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

Back to Last Lecture


```
X = 1
while FlipCoin(p) != HEADS:
   X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$\mathbb{E}[X] = \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) \quad \text{[Law of Total Expectation]}$$

$$= \mathbb{E}[X \mid H]\,p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) \quad \text{[The coin has bias } p\text{]}$$

$$= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) \quad \text{[If we get HEADS, we're done!]}$$

$$= p + \mathbb{E}[X+1](1-p) \quad [\star]$$

$$= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) \quad \text{[Linearity of Expectation]}$$

$$= 1 + (1-p)\mathbb{E}[X] \quad \text{[Simplifying]}$$

$$= \frac{1}{p} \quad \text{[Solve for } \mathbb{E}[X]\text{]}$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
  X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$\mathbb{E}[X] = \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right)$$ [Law of Total Expectation]

$$= \mathbb{E}[X \mid H]\,p + \mathbb{E}\left[X \mid \overline{H}\right](1-p)$$ [The coin has bias p]

$$= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p)$$ [If we get HEADS, we're done!]

$$= p + \mathbb{E}[X+1](1-p)$$ [$\star$]

$$= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p)$$ [Linearity of Expectation]

$$= 1 + (1-p)\mathbb{E}[X]$$ [Simplifying]

$$= \frac{1}{p}$$ [Solve for $\mathbb{E}[X]$]

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
    X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) && \text{[Law of Total Expectation]}\\
&= \mathbb{E}[X \mid H]\,p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[The coin has bias } p\text{]}\\
&= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[If we get HEADS, we're done!]}\\
&= p + \mathbb{E}[X+1](1-p) && [\star]\\
&= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) && \text{[Linearity of Expectation]}\\
&= 1 + (1-p)\mathbb{E}[X] && \text{[Simplifying]}\\
&= \frac{1}{p} && \text{[Solve for } \mathbb{E}[X]\text{]}
\end{aligned}$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
   X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up `HEADS`.

$$\mathbb{E}[X] = \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) \quad \text{[Law of Total Expectation]}$$

$$= \mathbb{E}[X \mid H]\, p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) \quad \text{[The coin has bias } p\text{]}$$

$$= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) \quad \text{[If we get HEADS, we're done!]}$$

$$= p + \mathbb{E}[X+1](1-p) \quad [\star]$$

$$= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) \quad \text{[Linearity of Expectation]}$$

$$= 1 + (1-p)\mathbb{E}[X] \quad \text{[Simplifying]}$$

$$= \frac{1}{p} \quad \text{[Solve for } \mathbb{E}[X]\text{]}$$

$\star$ – Since we know we got `TAILS` once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
   X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$\mathbb{E}[X] = \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) \quad \text{[Law of Total Expectation]}$$

$$= \mathbb{E}[X \mid H]\,p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) \quad \text{[The coin has bias } p\text{]}$$

$$= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) \quad \text{[If we get HEADS, we're done!]}$$

$$= p + \mathbb{E}[X+1](1-p) \quad [\star]$$

$$= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) \quad \text{[Linearity of Expectation]}$$

$$= 1 + (1-p)\mathbb{E}[X] \quad \text{[Simplifying]}$$

$$= \frac{1}{p} \quad \text{[Solve for } \mathbb{E}[X]\text{]}$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
    X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$\mathbb{E}[X] = \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) \quad \text{[Law of Total Expectation]}$$

$$= \mathbb{E}[X \mid H]\,p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) \quad \text{[The coin has bias } p\text{]}$$

$$= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) \quad \text{[If we get HEADS, we're done!]}$$

$$= p + \mathbb{E}[X+1](1-p) \quad [\star]$$

$$= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) \quad \text{[Linearity of Expectation]}$$

$$= 1 + (1-p)\mathbb{E}[X] \quad \text{[Simplifying]}$$

$$= \frac{1}{p} \quad \text{[Solve for } \mathbb{E}[X]\text{]}$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
    X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) && \text{[Law of Total Expectation]} \\
&= \mathbb{E}[X \mid H]\, p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[The coin has bias } p\text{]} \\
&= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[If we get HEADS, we're done!]} \\
&= p + \mathbb{E}[X+1](1-p) && [\star] \\
&= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) && \text{[Linearity of Expectation]} \\
&= 1 + (1-p)\mathbb{E}[X] && \text{[Simplifying]} \\
&= \frac{1}{p} && \text{[Solve for } \mathbb{E}[X]\text{]}
\end{aligned}$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
    X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$
\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) && \text{[Law of Total Expectation]} \\
&= \mathbb{E}[X \mid H]\, p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[The coin has bias } p\text{]} \\
&= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[If we get HEADS, we're done!]} \\
&= p + \mathbb{E}[X+1](1-p) && [\star] \\
&= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) && \text{[Linearity of Expectation]} \\
&= 1 + (1-p)\mathbb{E}[X] && \text{[Simplifying]} \\
&= \frac{1}{p} && \text{[Solve for } \mathbb{E}[X]\text{]}
\end{aligned}
$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.

```
X = 1
while FlipCoin(p) != HEADS:
   X++
```

Let X be the r.v. for how many loop iterations the code gets through.

Let H be the event that the next coin flip comes up HEADS.

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X \mid H]\Pr(H) + \mathbb{E}\left[X \mid \overline{H}\right]\Pr\left(\overline{H}\right) && \text{[Law of Total Expectation]}\\
&= \mathbb{E}[X \mid H]\,p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[The coin has bias } p\text{]}\\
&= p + \mathbb{E}\left[X \mid \overline{H}\right](1-p) && \text{[If we get HEADS, we're done!]}\\
&= p + \mathbb{E}[X+1](1-p) && [\star]\\
&= p + \mathbb{E}[X](1-p) + \mathbb{E}[1](1-p) && \text{[Linearity of Expectation]}\\
&= 1 + (1-p)\mathbb{E}[X] && \text{[Simplifying]}\\
&= \frac{1}{p} && \text{[Solve for } \mathbb{E}[X]\text{]}
\end{aligned}$$

$\star$ – Since we know we got TAILS once, we've flipped 1 coin. But ignoring that flip, we're at the same place we started at! So, $\mathbb{E}\left[X \mid \overline{H}\right] = \mathbb{E}[X+1]$.