CMPT 210: Probability and Computing

Lecture 16

Sharan Vaswani

March 12, 2024

Recap

Expectation/mean of a random variable R is denoted by $\mathbb{E}[R]$ and "summarizes" its distribution. Formally, $\mathbb{E}[R] := \sum_{\omega \in \mathcal{S}} \Pr[\omega]\, R[\omega]$

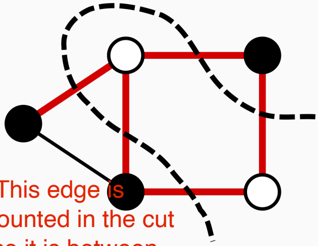
Alternate definition of expectation: $\mathbb{E}[R] = \sum_{x \in \text{Range}(R)} x \Pr[R = x]$.

Linearity of Expectation: For n random variables $R_1, R_2, \ldots, R_n$ and constants $a_1, a_2, \ldots, a_n$, $\mathbb{E}\left[\sum_{i=1}^{n} a_i R_i\right] = \sum_{i=1}^{n} a_i\, \mathbb{E}[R_i]$.

a_i is a constant which does not depend on any randomness.

Max Cut

Given a graph $G = (\mathcal{V}, \mathcal{E})$, partition the graph's vertices into two complementary sets $\mathcal{S}$ and $\mathcal{T}$, such that the number of edges between the set $\mathcal{S}$ and the set $\mathcal{T}$ is as large as possible.



This edge is
not counted in the cut
since it is between
nodes of the same color

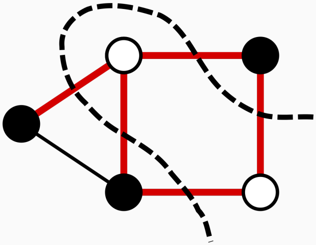
This comes up in the design of chips.

S U T = {set of all vertices}
S and T are disjoint.

If the black nodes are in S, and the white nodes are in T,

Max Cut

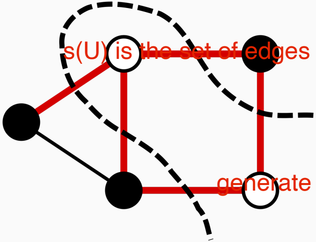
Given a graph $G = (\mathcal{V}, \mathcal{E})$, partition the graph's vertices into two complementary sets $\mathcal{S}$ and $\mathcal{T}$, such that the number of edges between the set $\mathcal{S}$ and the set $\mathcal{T}$ is as large as possible.



Max Cut has applications to VLSI circuit design.

Max Cut

Given a graph $G = (\mathcal{V}, \mathcal{E})$, partition the graph's vertices into two complementary sets $\mathcal{S}$ and $\mathcal{T}$, such that the number of edges between the set $\mathcal{S}$ and the set $\mathcal{T}$ is as large as possible.



s(U) is the set of edges that go inbetween black and white nodes.

Max Cut has applications to VLSI circuit design.

Possible algorithm:
generate all possible U's and then choose the one
with the largest cut.
This is an exaustive proof.

Equivalently, find a set $\mathcal{U} \subseteq \mathcal{V}$ of vertices that solve the following

$$\max_{\mathcal{U} \subseteq \mathcal{V}} |\delta(\mathcal{U})| \text{ where } \delta(\mathcal{U}) := \{(u, v) \in \mathcal{E} | u \in \mathcal{U} \text{ and } v \notin \mathcal{U}\}$$

Here, $\delta(\mathcal{U})$ is referred to as the "cut" corresponding to the set $\mathcal{U}$.

Max Cut

- Max Cut is NP-hard (Karp, 1972), meaning that there is no polynomial (in $|\mathcal{E}|$) time algorithm that solves Max Cut exactly. OPT is the size of the maximal cut.
- We want to find an approximate solution $\mathcal{U}$ such that, if OPT is the size of the optimal cut, then, $|\delta(\mathcal{U})| \geq \alpha\, OPT$ where $\alpha \in (0,1)$ is the multiplicative approximation factor.
- Randomized algorithm that guarantees an approximate solution with $\alpha = \frac{1}{2}$ with probability close to 1 (Erdos, 1967).
- Algorithm with $\alpha = 0.878$. (Goemans and Williamson, 1995).
- Under some technical conditions, no efficient algorithm has $\alpha > 0.878$ (Khot et al, 2004).

No algorithm can solve Max Cut in polynomial time.

The compromise made is that you output a cut that is bounded by the maximum cut.

If alpha = 1/2, then we will output a set that is at least 1/2 the optimal set cut.

We will generate a random algorithm that assures that alpha = 1/2.

Max Cut

- Max Cut is NP-hard (Karp, 1972), meaning that there is no polynomial (in $|\mathcal{E}|$) time algorithm that solves Max Cut exactly.
- We want to find an approximate solution $\mathcal{U}$ such that, if OPT is the size of the optimal cut, then, $|\delta(\mathcal{U})| \geq \alpha\, OPT$ where $\alpha \in (0, 1)$ is the multiplicative approximation factor.
- Randomized algorithm that guarantees an approximate solution with $\alpha = \frac{1}{2}$ with probability close to 1 (Erdos, 1967).
- Algorithm with $\alpha = 0.878$. (Goemans and Williamson, 1995).
- Under some technical conditions, no efficient algorithm has $\alpha > 0.878$ (Khot et al, 2004).

We will use Erdos' randomized algorithm and first prove the result in expectation. We wish to prove that for $\mathcal{U}$ returned by Erdos' algorithm,

$$\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} OPT$$

Max Cut

- Max Cut is NP-hard (Karp, 1972), meaning that there is no polynomial (in $|\mathcal{E}|$) time algorithm that solves Max Cut exactly.
- We want to find an approximate solution $\mathcal{U}$ such that, if OPT is the size of the optimal cut, then, $|\delta(\mathcal{U})| \geq \alpha\, OPT$ where $\alpha \in (0, 1)$ is the multiplicative approximation factor.
- Randomized algorithm that guarantees an approximate solution with $\alpha = \frac{1}{2}$ with probability close to 1 (Erdos, 1967).
- Algorithm with $\alpha = 0.878$. (Goemans and Williamson, 1995).
- Under some technical conditions, no efficient algorithm has $\alpha > 0.878$ (Khot et al, 2004).

We will use Erdos' randomized algorithm and first prove the result in expectation. We wish to prove that for $\mathcal{U}$ returned by Erdos' algorithm,

We do not know OPT, but know that it is the best possible value.

$$\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} OPT$$

O(size of vertices) since you make a decision at each node.

Algorithm: Select $\mathcal{U}$ to be a random subset of $\mathcal{V}$ i.e. for each vertex v, choose v to be in the set $\mathcal{U}$ independently with probability $\frac{1}{2}$ (do not even look at the edges!).

For each node, flip and coin and add it if the coin results in heads. Do not add it if you obtain tails otherwise.

Max Cut

Claim: For Erdos' algorithm, $\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} OPT$.

For each edge (v, e) in E, X_{v, e} is an indicator variable saying if {v, e} is in the cut. Zero otherwise.

E|s(U)| = # of edges that have different covered nodes

We can use linearity of expectation to push the expectation inside.

These events are mutually exclusive.

(v, e) is in the cut = (v is in U and e is not in U) or (v is not in U and e is in u)

Pr(v, e is in the cut) = pr(v in U, e not in U) + pr(v not in U, e in U)

Pr(v in U, e not in U) = Pr(v in U)*Pr(e not in U) = 1/2 * 1/2 = 1/4

Max Cut

Claim: For Erdos' algorithm, $\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} OPT$.

Proof: For each edge $(u, v) \in \mathcal{E}$, let $X_{u,v}$ be the indicator random variable equal to 1 iff the event $E_{u,v} = \{(u, v) \in \delta(\mathcal{U})\}$ happens.

Max Cut

Claim: For Erdos' algorithm, $\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} OPT$.

Proof: For each edge $(u, v) \in \mathcal{E}$, let $X_{u,v}$ be the indicator random variable equal to 1 iff the event $E_{u,v} = \{(u, v) \in \delta(\mathcal{U})\}$ happens.

$$\mathbb{E}[|\delta(\mathcal{U})|] = \mathbb{E}\left[\sum_{(u,v)\in\mathcal{E}} X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \mathbb{E}\left[X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \Pr[E_{u,v}]$$

(Linearity of expectation, and Expectation of indicator r.v's.)

Max Cut

Claim: For Erdos' algorithm, $\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} OPT$.

Proof: For each edge $(u, v) \in \mathcal{E}$, let $X_{u,v}$ be the indicator random variable equal to 1 iff the event $E_{u,v} = \{(u, v) \in \delta(\mathcal{U})\}$ happens.

$$\mathbb{E}[|\delta(\mathcal{U})|] = \mathbb{E}\left[\sum_{(u,v)\in\mathcal{E}} X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \mathbb{E}\left[X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \Pr[E_{u,v}]$$

(Linearity of expectation, and Expectation of indicator r.v's.)

$$\Pr[E_{u,v}] = \Pr[(u, v) \in \delta(\mathcal{U})] = \Pr\left[(u \in \mathcal{U} \cap v \notin \mathcal{U}) \cup (u \notin \mathcal{U} \cap v \in \mathcal{U})\right]$$

Max Cut

Claim: For Erdos' algorithm, $\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} OPT$.

Proof: For each edge $(u, v) \in \mathcal{E}$, let $X_{u,v}$ be the indicator random variable equal to 1 iff the event $E_{u,v} = \{(u, v) \in \delta(\mathcal{U})\}$ happens.

$$\mathbb{E}[|\delta(\mathcal{U})|] = \mathbb{E}\left[\sum_{(u,v)\in\mathcal{E}} X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \mathbb{E}\left[X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \Pr[E_{u,v}]$$

(Linearity of expectation, and Expectation of indicator r.v's.)

$$\Pr[E_{u,v}] = \Pr[(u, v) \in \delta(\mathcal{U})] = \Pr\left[(u \in \mathcal{U} \cap v \notin \mathcal{U}) \cup (u \notin \mathcal{U} \cap v \in \mathcal{U})\right]$$

$$= \Pr\left[(u \in \mathcal{U} \cap v \notin \mathcal{U})\right] + \Pr\left[(u \notin \mathcal{U} \cap v \in \mathcal{U})\right] \quad \text{(Union rule for mutually exclusive events)}$$

Max Cut

Claim: For Erdos' algorithm, $\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} OPT$.

Proof: For each edge $(u, v) \in \mathcal{E}$, let $X_{u,v}$ be the indicator random variable equal to 1 iff the event $E_{u,v} = \{(u, v) \in \delta(\mathcal{U})\}$ happens.

$$\mathbb{E}[|\delta(\mathcal{U})|] = \mathbb{E}\left[\sum_{(u,v)\in\mathcal{E}} X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \mathbb{E}\left[X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \Pr[E_{u,v}]$$

(Linearity of expectation, and Expectation of indicator r.v's.)

$$\Pr[E_{u,v}] = \Pr[(u, v) \in \delta(\mathcal{U})] = \Pr\left[(u \in \mathcal{U} \cap v \notin \mathcal{U}) \cup (u \notin \mathcal{U} \cap v \in \mathcal{U})\right]$$

$$= \Pr\left[(u \in \mathcal{U} \cap v \notin \mathcal{U})\right] + \Pr\left[(u \notin \mathcal{U} \cap v \in \mathcal{U})\right] \quad \text{(Union rule for mutually exclusive events)}$$

$$\Pr[E_{u,v}] = \Pr[u \in \mathcal{U}] \Pr[v \notin \mathcal{U}] + \Pr[u \notin \mathcal{U}] \Pr[v \in \mathcal{U}] = \frac{1}{2}\frac{1}{2} + \frac{1}{2}\frac{1}{2} = \frac{1}{2}.$$

(Independent events)

Max Cut

Claim: For Erdos' algorithm, $\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2}OPT$.

Proof: For each edge $(u, v) \in \mathcal{E}$, let $X_{u,v}$ be the indicator random variable equal to 1 iff the event $E_{u,v} = \{(u, v) \in \delta(\mathcal{U})\}$ happens.

$$\mathbb{E}[|\delta(\mathcal{U})|] = \mathbb{E}\left[\sum_{(u,v)\in\mathcal{E}} X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \mathbb{E}\left[X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \Pr[E_{u,v}]$$

(Linearity of expectation, and Expectation of indicator r.v's.)

$$\Pr[E_{u,v}] = \Pr[(u, v) \in \delta(\mathcal{U})] = \Pr\left[(u \in \mathcal{U} \cap v \notin \mathcal{U}) \cup (u \notin \mathcal{U} \cap v \in \mathcal{U})\right]$$

$$= \Pr\left[(u \in \mathcal{U} \cap v \notin \mathcal{U})\right] + \Pr\left[(u \notin \mathcal{U} \cap v \in \mathcal{U})\right]$$ (Union rule for mutually exclusive events)

$$\Pr[E_{u,v}] = \Pr[u \in \mathcal{U}] \Pr[v \notin \mathcal{U}] + \Pr[u \notin \mathcal{U}] \Pr[v \in \mathcal{U}] = \frac{1}{2}\frac{1}{2} + \frac{1}{2}\frac{1}{2} = \frac{1}{2}.$$

U can contain V, so the number of edges is an upper limit.

(Independent events)

$$\implies \mathbb{E}[|\delta(\mathcal{U})|] = \sum_{(u,v)\in\mathcal{E}} \Pr[E_{u,v}] = \frac{|\mathcal{E}|}{2} \geq \frac{\text{OPT}}{2}.$$

Here, E is the number of edges and alpha would be 1/2.

Max Cut

Claim: For Erdos' algorithm, $\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} OPT$.

Proof: For each edge $(u, v) \in \mathcal{E}$, let $X_{u,v}$ be the indicator random variable equal to 1 iff the event $E_{u,v} = \{(u, v) \in \delta(\mathcal{U})\}$ happens.

$$\mathbb{E}[|\delta(\mathcal{U})|] = \mathbb{E}\left[\sum_{(u,v)\in\mathcal{E}} X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \mathbb{E}\left[X_{u,v}\right] = \sum_{(u,v)\in\mathcal{E}} \Pr[E_{u,v}]$$

(Linearity of expectation, and Expectation of indicator r.v's.)

$$\Pr[E_{u,v}] = \Pr[(u, v) \in \delta(\mathcal{U})] = \Pr\left[(u \in \mathcal{U} \cap v \notin \mathcal{U}) \cup (u \notin \mathcal{U} \cap v \in \mathcal{U})\right]$$

$$= \Pr\left[(u \in \mathcal{U} \cap v \notin \mathcal{U})\right] + \Pr\left[(u \notin \mathcal{U} \cap v \in \mathcal{U})\right] \quad \text{(Union rule for mutually exclusive events)}$$

$$\Pr[E_{u,v}] = \Pr[u \in \mathcal{U}]\ \Pr[v \notin \mathcal{U}] + \Pr[u \notin \mathcal{U}]\ \Pr[v \in \mathcal{U}] = \frac{1}{2}\frac{1}{2} + \frac{1}{2}\frac{1}{2} = \frac{1}{2}.$$

(Independent events)

$$\implies \mathbb{E}[|\delta(\mathcal{U})|] = \sum_{(u,v)\in\mathcal{E}} \Pr[E_{u,v}] = \frac{|\mathcal{E}|}{2} \geq \frac{\text{OPT}}{2}.$$

Questions?

Conditional Expectation

Similar to probabilities, expectations can be conditioned on some event.

For random variable R, the expected value of R conditioned on an event A is given by:

$$\mathbb{E}[R|A] = \sum_{x \in \text{Range}(R)} x \Pr[R = x|A]$$

Conditional Expectation

Similar to probabilities, expectations can be conditioned on some event.

For random variable R, the expected value of R conditioned on an event A is given by:

$$\mathbb{E}[R|A] = \sum_{x \in \text{Range}(R)} x \Pr[R = x|A]$$

Q: If we throw a standard dice and define R to be the random variable equal to the number that comes up, what is the expected value of R given that the number is at most 4?

A subset of a uniform probability space is still a unifrom probability space?

E[Rl given number is at most four]

= 1 * Pr(R = 1 l at most four) + 2 * Pr(R = 2 l AMF) + 3 * Pr(R = 3 l AMF) + 4 * Pr(R = 4l AMF) + 5 * Pr(R = 5l AMF) + 6 * Pr(R = 6l AMF)

= 1 * 1/4 + 2 * 1/4 + 3 * 1/4 + 4 * 1/4 = 10/4 = 2.5

Conditional Expectation

Similar to probabilities, expectations can be conditioned on some event.

For random variable R, the expected value of R conditioned on an event A is given by:

$$\mathbb{E}[R|A] = \sum_{x \in \text{Range}(R)} x \Pr[R = x|A]$$

Q: If we throw a standard dice and define R to be the random variable equal to the number that comes up, what is the expected value of R given that the number is at most 4?

Let A be the event that the number is at most 4.
$\Pr[R = 1|A] = \frac{\Pr[(R=1) \cap A]}{\Pr[A]} = \frac{\Pr[R=1]}{\Pr[A]} = \frac{1/6}{4/6} = 1/4$.
$\Pr[R = 2|A] = \Pr[R = 3|A] = \Pr[R = 4|A] = \frac{1}{4}$ and $\Pr[R = 5|A] = \Pr[R = 6|A] = 0$.

$$\mathbb{E}[R|A] = \sum_{x \in \{1,2,3,4\}} x \Pr[R = x|A] = \frac{1}{4}[1 + 2 + 3 + 4] = \frac{5}{2}.$$

Conditional Expectation

Similar to probabilities, expectations can be conditioned on some event.

For random variable R, the expected value of R conditioned on an event A is given by:

$$\mathbb{E}[R|A] = \sum_{x \in \text{Range}(R)} x \Pr[R = x|A]$$

Q: If we throw a standard dice and define R to be the random variable equal to the number that comes up, what is the expected value of R given that the number is at most 4?

Let A be the event that the number is at most 4.
$\Pr[R = 1|A] = \frac{\Pr[(R=1) \cap A]}{\Pr[A]} = \frac{\Pr[R=1]}{\Pr[A]} = \frac{1/6}{4/6} = 1/4$.
$\Pr[R = 2|A] = \Pr[R = 3|A] = \Pr[R = 4|A] = \frac{1}{4}$ and $\Pr[R = 5|A] = \Pr[R = 6|A] = 0$.

$$\mathbb{E}[R|A] = \sum_{x \in \{1,2,3,4\}} x \Pr[R = x|A] = \frac{1}{4}[1 + 2 + 3 + 4] = \frac{5}{2}.$$

Q: What is the expected value of R given that the number is at least 4?

4 * 1/3 + 5 * 1/3 + 6 * 1/3 = 15/3 = 5

Law of Total Expectation

If R is a random variable $\mathcal{S} \to V$ and events $A_1, A_2, \ldots A_n$ form a partition of the sample space i.e. for all i, j, $A_i \cap A_j = \emptyset$ and $A_1 \cup A_2 \cup \ldots \cup A_n = \mathcal{S}$, then,

$$\mathbb{E}[R] = \sum_i \mathbb{E}[R|A_i] \Pr[A_i] \,.$$

Law of Total Expectation

If R is a random variable $\mathcal{S} \to V$ and events $A_1, A_2, \ldots A_n$ form a partition of the sample space i.e. for all i, j, $A_i \cap A_j = \emptyset$ and $A_1 \cup A_2 \cup \ldots \cup A_n = \mathcal{S}$, then,

$$\mathbb{E}[R] = \sum_i \mathbb{E}[R|A_i] \Pr[A_i] \,.$$

Proof :

$$\mathbb{E}[R] = \sum_{x \in \text{Range}(R)} x \Pr[R = x] = \sum_{x \in \text{Range}(R)} x \sum_i \Pr[R = x|A_i] \Pr[A_i]$$

(Law of total probability)

Law of Total Expectation

If R is a random variable $\mathcal{S} \to V$ and events $A_1, A_2, \ldots A_n$ form a partition of the sample space i.e. for all i, j, $A_i \cap A_j = \emptyset$ and $A_1 \cup A_2 \cup \ldots \cup A_n = \mathcal{S}$, then,

$$\mathbb{E}[R] = \sum_i \mathbb{E}[R|A_i] \Pr[A_i] .$$

Proof:

$$\mathbb{E}[R] = \sum_{x \in \text{Range}(R)} x \Pr[R = x] = \sum_{x \in \text{Range}(R)} x \sum_i \Pr[R = x|A_i] \Pr[A_i]$$

(Law of total probability)

$$= \sum_i \Pr[A_i] \sum_{x \in \text{Range}(R)} x \Pr[R = x|A_i]$$

Law of Total Expectation

If R is a random variable $\mathcal{S} \to V$ and events $A_1, A_2, \ldots A_n$ form a partition of the sample space i.e. for all i, j, $A_i \cap A_j = \emptyset$ and $A_1 \cup A_2 \cup \ldots \cup A_n = \mathcal{S}$, then,

$$\mathbb{E}[R] = \sum_i \mathbb{E}[R|A_i] \Pr[A_i] \,.$$

Proof:

$$\mathbb{E}[R] = \sum_{x \in \text{Range}(R)} x \Pr[R = x] = \sum_{x \in \text{Range}(R)} x \sum_i \Pr[R = x|A_i] \Pr[A_i]$$

(Law of total probability)

$$= \sum_i \Pr[A_i] \sum_{x \in \text{Range}(R)} x \Pr[R = x|A_i]$$

$$\implies \mathbb{E}[R] = \sum_i \Pr[A_i] \, \mathbb{E}[R|A_i].$$

Law of Total Expectation

If R is a random variable $\mathcal{S} \to V$ and events $A_1, A_2, \ldots A_n$ form a partition of the sample space i.e. for all i, j, $A_i \cap A_j = \emptyset$ and $A_1 \cup A_2 \cup \ldots \cup A_n = \mathcal{S}$, then,

$$\mathbb{E}[R] = \sum_i \mathbb{E}[R|A_i] \Pr[A_i] \,.$$

Proof:

$$\mathbb{E}[R] = \sum_{x \in \text{Range}(R)} x \Pr[R = x] = \sum_{x \in \text{Range}(R)} x \sum_i \Pr[R = x|A_i] \Pr[A_i]$$

(Law of total probability)

$$= \sum_i \Pr[A_i] \sum_{x \in \text{Range}(R)} x \Pr[R = x|A_i]$$

$$\implies \mathbb{E}[R] = \sum_i \Pr[A_i] \, \mathbb{E}[R|A_i].$$

Q: Suppose that 49.6% of the people in the world are male and the rest female. If the expected height of a randomly chosen male is 5 feet 11 inches, while the expected height of a randomly chosen female is 5 feet 5 inches, what is the expected height of a randomly chosen person?

Conditional Expectation - Examples

Q: Suppose that 49.6% of the people in the world are male and the rest female. If the expected height of a randomly chosen male is 5 feet 11 inches, while the expected height of a randomly chosen female is 5 feet 5 inches, what is the expected height of a randomly chosen person?

Define H to be the random variable equal to the height (in feet) of a randomly chosen person.
Define M to be the event that the person is male and F the event that the person is female.
We wish to compute $\mathbb{E}[H]$ and we know that $\mathbb{E}[H|M] = 5 + \frac{11}{12}$ and $\mathbb{E}[H|F] = 5 + \frac{5}{12}$.
$\Pr[M] = 0.496$ and $\Pr[F] = 1 - 0.496 = 0.504$.
Hence, $\mathbb{E}[H] = \mathbb{E}[H|M] \Pr[M] + \mathbb{E}[H|F] \Pr[F] = \frac{71}{12}(0.496) + \frac{65}{12}(0.504)$.

Questions?