Entropy, Randomness, and Information

Lecture 23 November 13, 2015

Part I

Entropy

"If only once - only once - no matter where, no matter before what audience - I could better the record of the great Rastelli and juggle with thirteen balls, instead of my usual twelve, I would feel that I had truly accomplished something for my country. But I am not getting any younger, and although I am still at the peak of my powers there are moments - why deny it? - when I begin to doubt and there is a time limit on all of us."

-Romain Gary, The talent scout.

Entropy: Definition

Definition

The **entropy** in bits of a discrete random variable X is

$$\mathbb{H}(X) = -\sum_x \Prig[X=xig] \lg \Prig[X=xig]\,.$$

Equivalently,
$$\mathbb{H}(X) = \mathrm{E}\Big[\lg rac{1}{\Pr[X]}\Big]$$
.

Entropy intuition...

Intuition...

 $\mathbb{H}(X)$ is the number of **fair** coin flips that one gets when getting the value of X.

Interpretation from last lecture...

Consider a (huge) string $S=s_1s_2\dots s_n$ formed by picking characters independently according to X. Then

$$|S| \, \mathbb{H}(X) = n \mathbb{H}(X)$$

is the minimum number of bits one needs to store the string S.

$$\mathbb{H}(X) = -\sum_x \Prig[X = xig] \lg \Prig[X = xig]$$

Definition

The **binary entropy** function $\mathbb{H}(p)$ for a random binary variable that is 1 with probability p, is $\mathbb{H}(p) = -p \lg p - (1-p) \lg (1-p)$. We define $\mathbb{H}(0) = \mathbb{H}(1) = 0$.

Q: How many truly random bits are there when given the result of flipping a single coin with probability p for heads?

Sariel (UIUC) New CS473 6 Fall 2015 6 / 30

$$\mathbb{H}(X) = -\sum_x \Pr \Big[X = x \Big] \lg \Pr \Big[X = x \Big]$$

Definition

The **binary entropy** function $\mathbb{H}(p)$ for a random binary variable that is 1 with probability p, is $\mathbb{H}(p) = -p \lg p - (1-p) \lg (1-p)$. We define $\mathbb{H}(0) = \mathbb{H}(1) = 0$.

Q: How many truly random bits are there when given the result of flipping a single coin with probability p for heads?

Sariel (UIUC) New CS473 6 Fall 2015 6 / 30

$$\mathbb{H}(X) = -\sum_x \Prig[X = xig] \lg \Prig[X = xig]$$

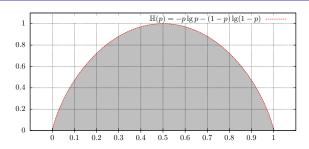
Definition

The **binary entropy** function $\mathbb{H}(p)$ for a random binary variable that is 1 with probability p, is $\mathbb{H}(p) = -p \lg p - (1-p) \lg (1-p)$. We define $\mathbb{H}(0) = \mathbb{H}(1) = 0$.

Q: How many truly random bits are there when given the result of flipping a single coin with probability p for heads?

Sariel (UIUC) New CS473 6 Fall 2015 6 / 30

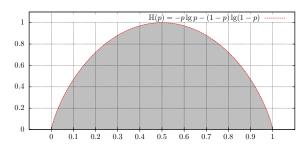
$$\mathbb{H}(p) = -p \lg p - (1-p) \lg (1-p)$$



- $lackbox{1}{} \mathbb{H}(p)$ is a concave symmetric around 1/2 on the interval [0,1].
- \bigcirc maximum at 1/2.
- **3** $\mathbb{H}(3/4) \approx 0.8113$ and $\mathbb{H}(7/8) \approx 0.5436$.
- \Longrightarrow coin that has 3/4 probably to be heads have higher amount of "randomness" in it than a coin that has probability 7/8 for heads.

Sariel (UIUC) New CS473 7 Fall 2015 7 / 30

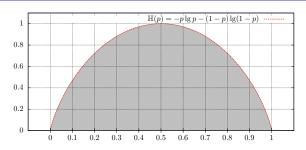
$$\mathbb{H}(p) = -p \lg p - (1-p) \lg (1-p)$$



- $lackbox{1}{} \mathbb{H}(p)$ is a concave symmetric around 1/2 on the interval [0,1].
- \bigcirc maximum at 1/2.
- **3** $\mathbb{H}(3/4) \approx 0.8113$ and $\mathbb{H}(7/8) \approx 0.5436$.
- \Longrightarrow coin that has 3/4 probably to be heads have higher amount of "randomness" in it than a coin that has probability 7/8 for heads.

Sariel (UIUC) New CS473 7 Fall 2015 7 / 30

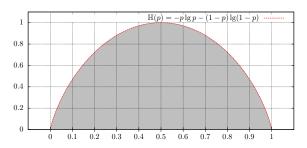
$$\mathbb{H}(p) = -p \lg p - (1-p) \lg (1-p)$$



- $lackbox{1}{} \mathbb{H}(p)$ is a concave symmetric around 1/2 on the interval [0,1].
- \bigcirc maximum at 1/2.
- **3** $\mathbb{H}(3/4) \approx 0.8113$ and $\mathbb{H}(7/8) \approx 0.5436$.
- \Longrightarrow coin that has 3/4 probably to be heads have higher amount of "randomness" in it than a coin that has probability 7/8 for heads.

Sariel (UIUC) New CS473 7 Fall 2015 7 / 30

$$\mathbb{H}(p) = -p \lg p - (1-p) \lg (1-p)$$



- $lackbox{1}{} \mathbb{H}(p)$ is a concave symmetric around 1/2 on the interval [0,1].
- \bigcirc maximum at 1/2.
- **3** $\mathbb{H}(3/4) \approx 0.8113$ and $\mathbb{H}(7/8) \approx 0.5436$.
- \implies coin that has 3/4 probably to be heads have higher amount of "randomness" in it than a coin that has probability 7/8 for heads.

- ② $\mathbb{H}'(p) = -\lg p + \lg(1-p) = \lg \frac{1-p}{p}$

- balanced coin has the largest amount of randomness in it.

- **3** $\mathbb{H}'(p) = -\lg p + \lg(1-p) = \lg \frac{1-p}{p}$

- balanced coin has the largest amount of randomness in it.

- **3** $\mathbb{H}'(p) = -\lg p + \lg(1-p) = \lg \frac{1-p}{p}$
- $\textcircled{1} \implies \mathbb{H}''(p) \leq 0$, for all $p \in (0,1)$, and the $\mathbb{H}(\cdot)$ is concave.
- $ext{ } ext{ } ext$

- **3** $\mathbb{H}'(p) = -\lg p + \lg(1-p) = \lg \frac{1-p}{p}$
- $\blacksquare \implies \mathbb{H}''(p) \leq 0$, for all $p \in (0,1)$, and the $\mathbb{H}(\cdot)$ is concave.
- balanced coin has the largest amount of randomness in it.

- **2** $\mathbb{H}'(p) = -\lg p + \lg(1-p) = \lg \frac{1-p}{p}$
- \blacksquare \Longrightarrow $\mathbb{H}''(p) \leq 0$, for all $p \in (0,1)$, and the $\mathbb{H}(\cdot)$ is concave.
- $ext{ } ext{ } ext$

- **3** $\mathbb{H}'(p) = -\lg p + \lg(1-p) = \lg \frac{1-p}{p}$
- \blacksquare \Longrightarrow $\mathbb{H}''(p) \leq 0$, for all $p \in (0,1)$, and the $\mathbb{H}(\cdot)$ is concave.
- balanced coin has the largest amount of randomness in it.

- $0 b_1, \ldots, b_n$: result of n coin flips...
- Prom a faulty coin!
- p: probability for head.
- We need fair bit coins!
- \bullet Convert $b_1, \ldots, b_n \implies b'_1, \ldots, b'_m$.
- **1** New bits must be truly random: Probability for head is 1/2.
- Q: How many truly random bits can we extract?

- $0 b_1, \ldots, b_n$: result of n coin flips...
- From a faulty coin!
- We need fair bit coins!
- $lacksquare{0}$ Convert $b_1,\ldots,b_n \implies b_1',\ldots,b_m'$
- **1** New bits must be truly random: Probability for head is 1/2.
- Q: How many truly random bits can we extract?

- $\mathbf{0}$ b_1, \ldots, b_n : result of n coin flips...
- From a faulty coin!
- We need fair bit coins!
- \bullet Convert $b_1,\ldots,b_n \implies b'_1,\ldots,b'_m$
- **1** New bits must be truly random: Probability for head is 1/2.
- Q: How many truly random bits can we extract?

- $\mathbf{0}$ b_1, \ldots, b_n : result of n coin flips...
- From a faulty coin!
- We need fair bit coins!
- $lacksquare{0}$ Convert $b_1,\ldots,b_n\implies b_1',\ldots,b_m'$
- **1** New bits must be truly random: Probability for head is 1/2.
- Q: How many truly random bits can we extract?

- b_1, \ldots, b_n : result of n coin flips...
- From a faulty coin!
- p: probability for head.
- We need fair bit coins!
- $\bullet \; \mathsf{Convert} \; b_1, \ldots, b_n \; \Longrightarrow \; b_1', \ldots, b_m'.$
- **1** New bits must be truly random: Probability for head is 1/2.
- Q: How many truly random bits can we extract?

- b_1, \ldots, b_n : result of n coin flips...
- From a faulty coin!
- p: probability for head.
- We need fair bit coins!
- $\bullet \; \mathsf{Convert} \; b_1, \ldots, b_n \; \Longrightarrow \; b_1', \ldots, b_m'.$
- **10** New bits must be truly random: Probability for head is 1/2.
- Q: How many truly random bits can we extract?

- b_1, \ldots, b_n : result of n coin flips...
- From a faulty coin!
- p: probability for head.
- We need fair bit coins!
- $\bullet \; \mathsf{Convert} \; b_1, \ldots, b_n \; \Longrightarrow \; b_1', \ldots, b_m'.$
- **10** New bits must be truly random: Probability for head is 1/2.
- Q: How many truly random bits can we extract?

Intuitively...

Squeezing good random bits out of bad random bits...

Question...

Given the result of n coin flips: b_1, \ldots, b_n from a faulty coin, with head with probability p, how many truly random bits can we extract?

If believe intuition about entropy, then this number should be $pprox n\mathbb{H}(p)$.

- lacksquareentropy of X is $\mathbb{H}(X) = -\sum_x \Prig[X = xig] \lg \Prig[X = xig].$
- Entropy of uniform variable.

Example

A random variable X that has probability 1/n to be i, for $i=1,\ldots,n$, has entropy $\mathbb{H}(X)=-\sum_{i=1}^n \frac{1}{n}\lg \frac{1}{n}=\lg n$.

- Entropy is oblivious to the exact values random variable can have.
- \implies random variables over -1,+1 with equal probability has the same entropy (i.e., 1) as a fair coin.

- lacksquareentropy of X is $\mathbb{H}(X) = -\sum_x \Prig[X = xig] \lg \Prig[X = xig].$
- Entropy of uniform variable..

Example

A random variable X that has probability 1/n to be i, for $i=1,\ldots,n$, has entropy $\mathbb{H}(X)=-\sum_{i=1}^n \frac{1}{n}\lg \frac{1}{n}=\lg n$.

- Entropy is oblivious to the exact values random variable can have.
- \Longrightarrow random variables over -1,+1 with equal probability has the same entropy (i.e., 1) as a fair coin.

- lacksquareentropy of X is $\mathbb{H}(X) = -\sum_x \Prig[X = xig] \lg \Prig[X = xig]$.
- Entropy of uniform variable..

Example

A random variable X that has probability 1/n to be i, for $i=1,\ldots,n$, has entropy $\mathbb{H}(X)=-\sum_{i=1}^n \frac{1}{n} \lg \frac{1}{n}=\lg n$.

- Entropy is oblivious to the exact values random variable can have.
- \implies random variables over -1,+1 with equal probability has the same entropy (i.e., 1) as a fair coin.

- lacksquareentropy of X is $\mathbb{H}(X) = -\sum_x \Pr\Big[X = x\Big] \lg \Pr\Big[X = x\Big].$
- Entropy of uniform variable..

Example

A random variable X that has probability 1/n to be i, for $i=1,\ldots,n$, has entropy $\mathbb{H}(X)=-\sum_{i=1}^n \frac{1}{n} \lg \frac{1}{n} = \lg n$.

- Entropy is oblivious to the exact values random variable can have.
- \implies random variables over -1,+1 with equal probability has the same entropy (i.e., 1) as a fair coin.

- lacksquareentropy of X is $\mathbb{H}(X) = -\sum_x \Prig[X = xig] \lg \Prig[X = xig].$
- Entropy of uniform variable..

Example

A random variable X that has probability 1/n to be i, for $i=1,\ldots,n$, has entropy $\mathbb{H}(X)=-\sum_{i=1}^n \frac{1}{n} \lg \frac{1}{n} = \lg n$.

- Entropy is oblivious to the exact values random variable can have.
- \Longrightarrow random variables over -1, +1 with equal probability has the same entropy (i.e., 1) as a fair coin.

Lemma: Entropy additive for independent variables

Lemma

Let X and Y be two independent random variables, and let Z be the random variable (X,Y). Then $\mathbb{H}(Z)=\mathbb{H}(X)+\mathbb{H}(Y)$.

Proof

In the following, summation are over all possible values that the variables can have. By the independence of $m{X}$ and $m{Y}$ we have

$$egin{aligned} \mathbb{H}(Z) &= \sum_{x,y} \Prig[(X,Y) = (x,y)ig] \lg rac{1}{\Prig[(X,Y) = (x,y)ig]} \ &= \sum_{x,y} \Prig[X = xig] \Prig[Y = yig] \lg rac{1}{\Prig[X = x] \Prig[Y = yig]} \ &= \sum_{x} \sum_{y} \Prig[X = xig] \Prig[Y = yig] \lg rac{1}{\Prig[X = xig]} \ &+ \sum_{x} \sum_{x} \Prig[X = xig] \Prig[Y = yig] \lg rac{1}{\Prig[Y = yig]} \end{aligned}$$

Sariel (UIUC) New CS473 13 Fall 2015 13 / 30

Proof continued

$$egin{aligned} \mathbb{H}(Z) &= \sum_x \sum_y \Pr[X=x] \Pr[Y=y] \lg rac{1}{\Pr[X=x]} \ &+ \sum_y \sum_x \Pr[X=x] \Pr[Y=y] \lg rac{1}{\Pr[Y=y]} \ &= \sum_x \Pr[X=x] \lg rac{1}{\Pr[X=x]} \ &+ \sum_y \Pr[Y=y] \lg rac{1}{\Pr[Y=y]} \ &= \mathbb{H}(X) + \mathbb{H}(Y). \end{aligned}$$

Sariel (UIUC) New CS473 14 Fall 2015 14 / 30

Bounding the binomial coefficient using entropy

Lemma

 $q \in [0,1]$ nq is integer in the range [0,n]. Then

$$\frac{2^{n\mathbb{H}(q)}}{n+1} \leq \binom{n}{nq} \leq 2^{n\mathbb{H}(q)}.$$

Proof

Holds if q=0 or q=1, so assume 0 < q < 1. We have

$$\binom{n}{nq}q^{nq}(1-q)^{n-nq} \le (q+(1-q))^n = 1$$

We also have:

$$q^{-nq}(1-q)^{-(1-q)n}=2^{n\;(-q\lg q-(1-q)\lg(1-q))}=2^{n\mathbb{H}(q)}$$
 , we have

$$\binom{n}{nq} \le q^{-nq} (1-q)^{-(1-q)n} = 2^{n\mathbb{H}(q)}.$$

Sariel (UIUC) New CS473 16 Fall 2015 16 / 30

Proof

Holds if q=0 or q=1, so assume 0 < q < 1. We have

$$\binom{n}{nq}q^{nq}(1-q)^{n-nq} \le (q+(1-q))^n = 1.$$

We also have:

$$q^{-nq}(1-q)^{-(1-q)n}=2^{n\;(-q\lg q-(1-q)\lg(1-q))}=2^{n\mathbb{H}(q)}$$
 , we have

$$inom{n}{nq} \le q^{-nq} (1-q)^{-(1-q)n} = 2^{n\mathbb{H}(q)}.$$

Sariel (UIUC) New CS473 16 Fall 2015 16 / 30

Proof

Holds if q=0 or q=1, so assume 0 < q < 1. We have

$$\binom{n}{nq}q^{nq}(1-q)^{n-nq} \le (q+(1-q))^n = 1.$$

We also have:

$$q^{-nq}(1-q)^{-(1-q)n}=2^{n\,(-q\,\lg q-(1-q)\,\lg(1-q))}=2^{n\mathbb{H}(q)}$$
 , we have

$$inom{n}{nq} \leq q^{-nq} (1-q)^{-(1-q)n} = 2^{n\mathbb{H}(q)}.$$

Sariel (UIUC) New CS473 16 Fall 2015 16 / 30

Proof

Holds if q=0 or q=1, so assume 0 < q < 1. We have

$$\binom{n}{nq}q^{nq}(1-q)^{n-nq} \le (q+(1-q))^n = 1.$$

We also have:

$$q^{-nq}(1-q)^{-(1-q)n}=2^{n\ (-q\lg q-(1-q)\lg(1-q))}=2^{n\mathbb{H}(q)}$$
 , we have

$$egin{pmatrix} n \ nq \end{pmatrix} \leq q^{-nq} (1-q)^{-(1-q)n} = 2^{n\mathbb{H}(q)}.$$

Sariel (UIUC) New CS473 16 Fall 2015 16 / 30

- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- $\bigcirc \sum_{i=0}^{n} \binom{n}{i} q^{i} (1-q)^{n-i} = \sum_{i=0}^{n} \mu(i).$
- ullet Claim: $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$ largest term in $\sum_{k=0}^n \mu(k)=1.$
- lacksquare sign of $\Delta_k=$ size of last term...
- $\begin{aligned} & \mathbf{o} \operatorname{sign}(\Delta_k) = \operatorname{sign}\left(1 \frac{(n-k)q}{(k+1)(1-q)}\right) \\ & = \operatorname{sign}\left(\frac{(k+1)(1-q) (n-k)q}{(k+1)(1-q)}\right) \end{aligned}$

- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- $igo \sum_{i=0}^n inom{n}{i} q^i (1-q)^{n-i} = \sum_{i=0}^n \mu(i).$
- ullet Claim: $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$ largest term in $\sum_{k=0}^n \mu(k)=1.$
- ullet sign of $\Delta_k=$ size of last term...
- $\begin{aligned} & \mathbf{o} \operatorname{sign}(\Delta_k) = \operatorname{sign}\left(1 \frac{(n-k)q}{(k+1)(1-q)}\right) \\ & = \operatorname{sign}\left(\frac{(k+1)(1-q) (n-k)q}{(k+1)(1-q)}\right) \end{aligned}$

- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- ullet Claim: $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$ largest term in $\sum_{k=0}^n \mu(k)=1.$
- ullet sign of $\Delta_k=$ size of last term...
- $\begin{aligned} & \mathbf{o} \operatorname{sign}(\Delta_k) = \operatorname{sign}\left(1 \frac{(n-k)q}{(k+1)(1-q)}\right) \\ & = \operatorname{sign}\left(\frac{(k+1)(1-q) (n-k)q}{(k+1)(1-q)}\right) \end{aligned}$

- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- ullet Claim: $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$ largest term in $\sum_{k=0}^n \mu(k)=1.$
- $oldsymbol{\Phi} \ \Delta_k = \mu(k) \mu(k+1) = inom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q} \Big),$
- ullet sign of $\Delta_k =$ size of last term...
- $\begin{aligned} & \mathbf{o} \operatorname{sign}(\Delta_k) = \operatorname{sign}\left(1 \frac{(n-k)q}{(k+1)(1-q)}\right) \\ & = \operatorname{sign}\left(\frac{(k+1)(1-q) (n-k)q}{(k+1)(1-q)}\right) \end{aligned}$

- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- ullet Claim: $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$ largest term in $\sum_{k=0}^n \mu(k)=1.$
- $oldsymbol{\Delta}_k = \mu(k) \mu(k+1) = inom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q} \Big),$
- **5** sign of Δ_k = size of last term...

- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- ullet Claim: $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$ largest term in $\sum_{k=0}^n \mu(k)=1.$
- $lack \Delta_k = \mu(k) \mu(k+1) = inom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q} \Big),$
- **5** sign of Δ_k = size of last term...
- $\begin{array}{l} \bullet \ \operatorname{sign}(\Delta_k) = \operatorname{sign}\Bigl(1 \frac{(n-k)q}{(k+1)(1-q)}\Bigr) \\ = \operatorname{sign}\Bigl(\frac{(k+1)(1-q) (n-k)q}{(k+1)(1-q)}\Bigr) \end{array}$

- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- 3 Claim: $\mu(nq)=\binom{n}{nq}q^{nq}(1-q)^{n-nq}$ largest term in $\sum_{k=0}^n \mu(k)=1$.
- $oldsymbol{\Delta}_k = \mu(k) \mu(k+1) = inom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q} \Big),$
- **5** sign of Δ_k = size of last term...

- $egin{array}{ll} igotimes \Delta_k \geq 0 & ext{when } k \geq nq+q-1 \ \Delta_k < 0 & ext{otherwise}. \end{array}$
- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- $lack 0 \ \mu(k) < \mu(k+1)$, for k < nq , and $\mu(k) \geq \mu(k+1)$ for $k \geq nq$.
- ullet $\mu(nq)$ larger than the average in sum.

Sariel (UIUC) New CS473 18 Fall 2015 18 / 30

- $igotimes \Delta_k \geq 0$ when $k \geq nq+q-1$ $\Delta_k < 0$ otherwise.
- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- $lack 0 \ \mu(k) < \mu(k+1), ext{ for } k < nq, ext{ and } \mu(k) \geq \mu(k+1) ext{ for } k \geq nq.$
- ullet $\mu(nq)$ larger than the average in sum.

- $egin{aligned} igotimes \Delta_k \geq 0 & ext{when } k \geq nq+q-1 \ \Delta_k < 0 & ext{otherwise}. \end{aligned}$
- \bullet $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- $lack 0 \ \mu(k) < \mu(k+1)$, for k < nq , and $\mu(k) \geq \mu(k+1)$ for $k \geq nq$.
- $\Longrightarrow \ \mu(nq) \$ is the largest term in $\sum_{k=0}^n \mu(k) = 1.$
- ullet $\mu(nq)$ larger than the average in sum.

- $egin{aligned} igotimes \Delta_k \geq 0 & ext{when } k \geq nq+q-1 \ \Delta_k < 0 & ext{otherwise}. \end{aligned}$
- \bullet $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- $\Longrightarrow \ \mu(nq) \$ is the largest term in $\sum_{k=0}^n \mu(k) = 1.$
- ullet $\mu(nq)$ larger than the average in sum.

- $egin{aligned} igotimes \Delta_k \geq 0 & ext{when } k \geq nq+q-1 \ \Delta_k < 0 & ext{otherwise}. \end{aligned}$
- \bullet $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- lacksquare \Rightarrow $\mu(nq)$ is the largest term in $\sum_{k=0}^n \mu(k) = 1$.
- ullet $\mu(nq)$ larger than the average in sum.

- $egin{aligned} igotimes \Delta_k & \geq 0 ext{ when } k \geq nq+q-1 \ \Delta_k & < 0 ext{ otherwise}. \end{aligned}$
- \bullet $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- lacksquare \Rightarrow $\mu(nq)$ is the largest term in $\sum_{k=0}^n \mu(k) = 1$.
- ullet $\mu(nq)$ larger than the average in sum.

- $egin{aligned} igotimes \Delta_k \geq 0 & ext{when } k \geq nq+q-1 \ \Delta_k < 0 & ext{otherwise}. \end{aligned}$
- \bullet $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- lacksquare \Rightarrow $\mu(nq)$ is the largest term in $\sum_{k=0}^n \mu(k) = 1$.
- ullet $\mu(nq)$ larger than the average in sum.

- $egin{aligned} igotimes \Delta_k & \geq 0 ext{ when } k \geq nq+q-1 \ \Delta_k & < 0 ext{ otherwise}. \end{aligned}$
- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- lacksquare \Rightarrow $\mu(nq)$ is the largest term in $\sum_{k=0}^n \mu(k) = 1$.
- ullet $\mu(nq)$ larger than the average in sum.
- $igotimes ig(rac{n}{nq} ig) \geq rac{1}{n+1} q^{-nq} (1-q)^{-(n-nq)} = rac{1}{n+1} 2^{n\mathbb{H}(q)}.$



Generalization...

Corollary

We have:

(i)
$$q \in [0, 1/2] \Rightarrow \binom{n}{\lfloor nq \rfloor} \le 2^{n\mathbb{H}(q)}$$
.

(ii)
$$q \in [1/2,1] \binom{n}{\lceil nq \rceil} \leq 2^{n\mathbb{H}(q)}$$
.

(iii)
$$q \in [1/2,1] \Rightarrow rac{2^{n\mathbb{H}(q)}}{n+1} \leq inom{n}{\lfloor nq \rfloor}.$$

(iv)
$$q \in [0,1/2] \Rightarrow rac{2^{n\mathbb{H}(q)}}{n+1} \leq {n \choose \lceil nq \rceil}$$
.

Proof is straightforward but tedious.

- lacksquare Proved that $\binom{n}{nq} pprox 2^{n\mathbb{H}(q)}$.
- Estimate is loose.
- Sanity check...
 - (I) A sequence of n bits generated by coin with probability q for head.
 - (II) By Chernoff inequality... roughly nq heads in this sequence.
 - (III) Generated sequence Y belongs to $\binom{n}{nq} pprox 2^{n\mathbb{H}(q)}$ possible sequences .
 - (IV) ...of similar probability.
 - $(V) \implies \mathbb{H}(Y) = n\mathbb{H}(q) \approx \lg \binom{n}{nq}.$

- lacksquare Proved that $\binom{n}{nq} pprox 2^{n\mathbb{H}(q)}$.
- Estimate is loose.
- Sanity check...
 - (I) A sequence of n bits generated by coin with probability q for head.
 - (II) By Chernoff inequality... roughly $\boldsymbol{n}\boldsymbol{q}$ heads in this sequence.
 - (III) Generated sequence Y belongs to $\binom{n}{nq} pprox 2^{n\mathbb{H}(q)}$ possible sequences .
 - (IV) ...of similar probability.
 - $(V) \implies \mathbb{H}(Y) = n\mathbb{H}(q) \approx \lg \binom{n}{nq}.$

- lacksquare Proved that $\binom{n}{nq} pprox 2^{n\mathbb{H}(q)}$.
- Estimate is loose.
- Sanity check...
 - (I) A sequence of n bits generated by coin with probability q for head.
 - (II) By Chernoff inequality... roughly nq heads in this sequence.
 - (III) Generated sequence Y belongs to $\binom{n}{nq} pprox 2^{n\mathbb{H}(q)}$ possible sequences .
 - (IV) ...of similar probability.
 - $(V) \implies \mathbb{H}(Y) = n\mathbb{H}(q) \approx \lg \binom{n}{nq}.$

- lacksquare Proved that $\binom{n}{nq} pprox 2^{n\mathbb{H}(q)}$.
- Estimate is loose.
- Sanity check...
 - (I) A sequence of n bits generated by coin with probability q for head.
 - (II) By Chernoff inequality... roughly nq heads in this sequence.
 - (III) Generated sequence Y belongs to $\binom{n}{nq} pprox 2^{n\mathbb{H}(q)}$ possible sequences .
 - (IV) ...of similar probability.
 - $(\mathsf{V}) \implies \mathbb{H}(Y) = n\mathbb{H}(q) pprox \lg inom{n}{nq}.$

- lacksquare Proved that $\binom{n}{nq} pprox 2^{n\mathbb{H}(q)}$.
- 2 Estimate is loose.
- Sanity check...
 - (I) A sequence of n bits generated by coin with probability q for head.
 - (II) By Chernoff inequality... roughly nq heads in this sequence.
 - (III) Generated sequence Y belongs to $\binom{n}{nq} pprox 2^{n\mathbb{H}(q)}$ possible sequences .
 - (IV) ...of similar probability.
 - $(\mathsf{V}) \implies \mathbb{H}(Y) = n\mathbb{H}(q) \approx \lg \binom{n}{nq}.$

Just one bit...

question

Given a coin C with:

p: Probability for head.

q = 1 - p: Probability for tail.

Q: How to get **one** true random bit, by flipping C.

Describe an algorithm!

Entropy can be interpreted as the amount of unbiased random coin flips can be extracted from a random variable.

Definition

An extraction function \mathbf{Ext} takes as input the value of a random variable \boldsymbol{X} and outputs a sequence of bits \boldsymbol{y} , such that

 $\Pr\Bigl[\mathsf{Ext}(X) = y \ \Big| \ |y| = k \Bigr] = \frac{1}{2^k}$, whenever $\Pr[|y| = k] > 0$, where |y| denotes the length of y.

- **1** X: uniform random integer variable out of $0, \ldots, 7$.
- 2 $\mathsf{Ext}(X)$: binary representation of x.
- Open Subtle: all extracted seqs of same len have same probability
- Another example of extraction scheme:
 - **1** X: uniform random integer variable $0, \ldots, 11$.
 - 2 Ext(x): output the binary representation for x if $0 \le x \le 7$

 - Idea... Output binary representation of x-8 as a two bit number.
- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- **1** X: uniform random integer variable out of $0, \ldots, 7$.
- **2** Ext(X): binary representation of x.
- Open Def. subtle: all extracted seqs of same len have same probability
- 4 Another example of extraction scheme:
 - **1** X: uniform random integer variable $0, \ldots, 11$.
 - 2 Ext(x): output the binary representation for x if $0 \le x \le 7$

 - Idea... Output binary representation of x-8 as a two bit number.
- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- **1** X: uniform random integer variable out of $0, \ldots, 7$.
- **2** Ext(X): binary representation of x.
- Oef. subtle: all extracted seqs of same len have same probability.
- 4 Another example of extraction scheme:
 - **1** X: uniform random integer variable $0, \ldots, 11$.
 - 2 Ext(x): output the binary representation for x if $0 \le x \le 7$

 - ullet Idea... Output binary representation of x-8 as a two bit number.
- $egin{aligned} \mathsf{A} ext{ valid extractor...} \ & \Pr \Big[\mathsf{Ext}(X) = 00 \ \Big| \ |\mathsf{Ext}(X)| = 2 \Big] = rac{1}{4}, \end{aligned}$

- **1** X: uniform random integer variable out of $0, \ldots, 7$.
- **2** Ext(X): binary representation of x.
- Oef. subtle: all extracted seqs of same len have same probability.
- 4 Another example of extraction scheme:
 - **1** X: uniform random integer variable $0, \ldots, 11$.
 - **2** Ext(x): output the binary representation for x if $0 \le x \le 7$

 - ullet Idea... Output binary representation of x-8 as a two bit number.
- $egin{aligned} \mathsf{A} ext{ valid extractor...} \ & \Pr \Big[\mathsf{Ext}(X) = 00 \ \Big| \ |\mathsf{Ext}(X)| = 2 \Big] = rac{1}{4}, \end{aligned}$

- **1** X: uniform random integer variable out of $0, \ldots, 7$.
- **2** Ext(X): binary representation of x.
- Oef. subtle: all extracted seqs of same len have same probability.
- 4 Another example of extraction scheme:
 - **1** X: uniform random integer variable $0, \ldots, 11$.
 - **2** Ext(x): output the binary representation for x if $0 \le x \le 7$.

 - ullet Idea... Output binary representation of x-8 as a two bit number.
- $egin{aligned} egin{aligned} \mathsf{A} ext{ valid extractor...} \ & \Pr \Big[\mathsf{Ext}(X) = 00 \ \Big| \ |\mathsf{Ext}(X)| = 2 \Big] = rac{1}{4}, \end{aligned}$

- **1** X: uniform random integer variable out of $0, \ldots, 7$.
- **2** Ext(X): binary representation of x.
- Oef. subtle: all extracted seqs of same len have same probability.
- 4 Another example of extraction scheme:
 - **1** X: uniform random integer variable $0, \ldots, 11$.
 - **2** Ext(x): output the binary representation for x if $0 \le x \le 7$.

 - ullet Idea... Output binary representation of x-8 as a two bit number.
- $egin{aligned} egin{aligned} \mathsf{A} \ \mathsf{valid} \ \mathsf{extractor...} \ \mathsf{Pr}\Big[\mathsf{Ext}(X) = 00 \ \Big| \ |\mathsf{Ext}(X)| = 2 \Big] = rac{1}{4}, \end{aligned}$

- **1 X**: uniform random integer variable out of $0, \ldots, 7$.
- **2** Ext(X): binary representation of x.
- Oef. subtle: all extracted seqs of same len have same probability.
- 4 Another example of extraction scheme:
 - **1** X: uniform random integer variable $0, \ldots, 11$.
 - **2** Ext(x): output the binary representation for x if $0 \le x \le 7$.

 - Idea... Output binary representation of x-8 as a two bit number.
- $egin{aligned} egin{aligned} \mathsf{A} ext{ valid extractor...} \ & \Pr \Big[\mathsf{Ext}(X) = 00 \ \Big| \ |\mathsf{Ext}(X)| = 2 \Big] = rac{1}{4}, \end{aligned}$

- **1** X: uniform random integer variable out of $0, \ldots, 7$.
- **2** Ext(X): binary representation of x.
- Oef. subtle: all extracted seqs of same len have same probability.
- 4 Another example of extraction scheme:
 - **1** X: uniform random integer variable $0, \ldots, 11$.
 - **2** Ext(x): output the binary representation for x if $0 \le x \le 7$.
 - 3 If x is between 8 and 11?
 - Idea... Output binary representation of x-8 as a two bit number.
- A valid extractor...

$$\Pr\Bigl[\mathsf{Ext}(X) = 00 \ \Big| \ |\mathsf{Ext}(X)| = 2 \Bigr] = rac{1}{4},$$

Technical lemma

The following is obvious, but we provide a proof anyway.

Lemma

Let x/y be a faction, such that x/y < 1. Then, for any i, we have x/y < (x+i)/(y+i).

Proof.

We need to prove that x(y+i)-(x+i)y<0. The left size is equal to i(x-y), but since y>x (as x/y<1), this quantity is negative, as required.

A uniform variable extractor...

Theorem

- **1** X: random variable chosen uniformly at random from $\{0, \ldots, m-1\}$.
- ullet Then there is an extraction function for X.
 - outputs on average at least

$$\lfloor \lg m
floor - 1 = \lfloor \mathbb{H}(X)
floor - 1$$

independent and unbiased bits.

A uniform variable extractor...

Theorem

- **1** X: random variable chosen uniformly at random from $\{0,\ldots,m-1\}$.
- Then there is an extraction function for X:
 - outputs on average at least

$$\lfloor \lg m
floor - 1 = \lfloor \mathbb{H}(X)
floor - 1$$

independent and unbiased bits.

A uniform variable extractor...

Theorem

- **1** X: random variable chosen uniformly at random from $\{0,\ldots,m-1\}$.
- ② Then there is an extraction function for X:
 - outputs on average at least

$$\lfloor \lg m
floor - 1 = \lfloor \mathbb{H}(X)
floor - 1$$

independent and unbiased bits.

- $oldsymbol{0}$ m: A sum of unique powers of $oldsymbol{2}$, namely $m=\sum_i a_i 2^i$, where $a_i \in \{0,1\}.$
- ② Example:
- ullet decomposed $\{0,\ldots,m-1\}$ into disjoint union of blocks sizes are powers of 2.
- ullet If x is in block 2^k , output its relative location in the block in binary representation.
- Example: x = 10: then falls into block 2^2 ... x relative location is 2. Output 2 written using two bits Output: "10".

Sariel (UIUC) New CS473 26 Fall 2015 26 / 30

 $m{0}$ $m{m}$: A sum of unique powers of $m{2}$, namely $m{m} = \sum_i a_i m{2}^i$, where $a_i \in \{0,1\}$.

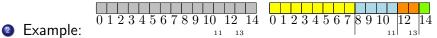


- ullet decomposed $\{0,\ldots,m-1\}$ into disjoint union of bloc

 - Example: x = 10: then falls into block 2^2 ... x relative location is 2. Output 2 written using two bits. Output: "10".

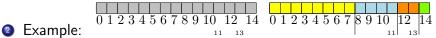
Sariel (UIUC) New CS473 26 Fall 2015 26 / 30

 $oldsymbol{0}$ m: A sum of unique powers of $oldsymbol{2}$, namely $m=\sum_i a_i 2^i$, where $a_i \in \{0,1\}.$



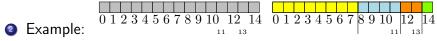
- **3** decomposed $\{0, \ldots, m-1\}$ into disjoint union of blocks sizes are powers of **2**.
- If x is in block 2^k , output its relative location in the block in binary representation.
- Example: x = 10: then falls into block 2²... x relative location is 2. Output 2 written using two bits, Output: "10".

 $oldsymbol{0}$ m: A sum of unique powers of $oldsymbol{2}$, namely $m=\sum_i a_i 2^i$, where $a_i \in \{0,1\}.$



- **1** decomposed $\{0, \ldots, m-1\}$ into disjoint union of blocks sizes are powers of 2.
- **1** If x is in block 2^k , output its relative location in the block in binary representation.
- Example: x = 10: then falls into block 2²... x relative location is 2. Output 2 written using two bits, Output: "10".

 $oldsymbol{0}$ m: A sum of unique powers of $oldsymbol{2}$, namely $m=\sum_i a_i 2^i$, where $a_i \in \{0,1\}$.

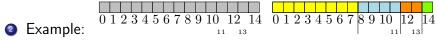


- **3** decomposed $\{0,\ldots,m-1\}$ into disjoint union of blocks sizes are powers of **2**.
- ullet If x is in block 2^k , output its relative location in the block in binary representation.



then falls into block 2^2 ... x relative location is 2. Output 2 written using two bits, Output: "10".

 $oldsymbol{0}$ m: A sum of unique powers of $oldsymbol{2}$, namely $m=\sum_i a_i 2^i$, where $a_i \in \{0,1\}.$

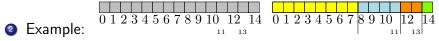


- ullet If x is in block 2^k , output its relative location in the block in binary representation.



- then falls into block 2^2 ...
 - *x* relative location is 2. Output 2 written using two bits, Output: "10".

 $oldsymbol{0}$ m: A sum of unique powers of $oldsymbol{2}$, namely $m=\sum_i a_i 2^i$, where $a_i \in \{0,1\}.$



- **3** decomposed $\{0,\ldots,m-1\}$ into disjoint union of blocks sizes are powers of **2**.
- ullet If x is in block 2^k , output its relative location in the block in binary representation.



then falls into block 2^2 ... x relative location is 2. Output 2 written using two bits,

Output: "10"

 $oldsymbol{0}$ m: A sum of unique powers of $oldsymbol{2}$, namely $m=\sum_i a_i 2^i$, where $a_i \in \{0,1\}.$



- ullet If x is in block 2^k , output its relative location in the block in binary representation.



then falls into block 2^2 ... x relative location is 2. Output 2 written using two bits, Output: "10".

- Valid extractor...
- 2 Theorem holds if m is a power of two. Only one block.
- \bigcirc m not a power of 2...
- ullet X falls in block of size 2^k : then output k complete random bits..

 \dots entropy is k

- ullet Let $2^k < m < 2^{k+1}$ biggest block.
- $egin{aligned} \mathbf{u} &= \left\lfloor \lg(m-2^k)
 ight
 floor < k. \end{aligned}$ There must be a block of size u in the decomposition of m.
- Largest two blocks...
- \bigcirc Y: random variable = number of bits output by extractor.

- Valid extractor...
- ② Theorem holds if m is a power of two. Only one block.
- \bigcirc m not a power of 2...
- ullet X falls in block of size 2^k : then output k complete random bits..
 - \dots entropy is k
- ullet Let $2^k < m < 2^{k+1}$ biggest block.
- $egin{aligned} \mathbf{u} &= \left\lfloor \lg(m-2^k)
 ight
 floor < k. \end{aligned}$ There must be a block of size u in the decomposition of m.
- \bigcirc two blocks in decomposition of m: sizes 2^k and 2^u .
- a Largest two blocks...
- \bigcirc Y: random variable = number of bits output by extractor.

- Valid extractor...
- ② Theorem holds if m is a power of two. Only one block.
- \bigcirc m not a power of 2...
- ullet X falls in block of size 2^k : then output k complete random bits..
 - \dots entropy is k
- ullet Let $2^k < m < 2^{k+1}$ biggest block.
- $oxed{0} \ u = \left\lfloor \lg(m-2^k)
 ight
 floor < k.$ There must be a block of size u in the decomposition of m.
- Largest two blocks...

- Valid extractor...
- ② Theorem holds if m is a power of two. Only one block.
- $oldsymbol{3}$ m not a power of $oldsymbol{2}$...
- ullet X falls in block of size 2^k : then output k complete random bits..
 - \dots entropy is k
- ullet Let $2^k < m < 2^{k+1}$ biggest block.
- $oxed{0} \ u = \left\lfloor \lg(m-2^k)
 ight
 floor < k.$ There must be a block of size u in the decomposition of m.
- \bigcirc two blocks in decomposition of m: sizes 2^k and 2^u .
- a Largest two blocks...

- Valid extractor...
- $oldsymbol{0}$ m not a power of $oldsymbol{2}$...
- **4** It is a block of size 2^k : then output k complete random bits..
 - ... entropy is k.
- ullet Let $2^k < m < 2^{k+1}$ biggest block.
- $ullet u = \left\lfloor \lg(m-2^k)
 ight
 floor < k.$ There must be a block of size u in the decomposition of m.
- Largest two blocks...

- Valid extractor...
- $oldsymbol{0}$ m not a power of $oldsymbol{2}$...
- **3 a** X falls in block of size 2^k : then output k complete random bits..
 - ... entropy is k.
- ullet Let $2^k < m < 2^{k+1}$ biggest block.
- $oxed{0} \ u = \left\lfloor \lg(m-2^k)
 ight
 floor < k.$ There must be a block of size u in the decomposition of m.
- Largest two blocks...

- Valid extractor...
- $oldsymbol{0}$ m not a power of $oldsymbol{2}$...
- **3 a** X falls in block of size 2^k : then output k complete random bits..
 - ... entropy is k.
- **1** Let $2^k < m < 2^{k+1}$ biggest block.
- $oldsymbol{0} u = \left\lfloor \lg(m-2^k) \right\rfloor < k.$ There must be a block of size u in the decomposition of m.
- @ two blocks in decomposition of m: sizes 2^k and 2^u .
- a Largest two blocks...
- \bigcirc Y: random variable = number of bits output by extractor.

- Valid extractor...
- $oldsymbol{0}$ m not a power of $oldsymbol{2}$...
- **3** X falls in block of size 2^k : then output k complete random bits..
 - ... entropy is k.
- **1** Let $2^k < m < 2^{k+1}$ biggest block.
- $\mathbf{0} \ \ u = \lfloor \lg(m-2^k) \rfloor < k.$ There must be a block of size u in the decomposition of m.
- lacksquare two blocks in decomposition of m: sizes 2^k and 2^u .
- a Largest two blocks...
- ${ t @} \; Y$: random variable = number of bits output by extractor.

- Valid extractor...
- $oldsymbol{0}$ m not a power of $oldsymbol{2}$...
- **3 a** X falls in block of size 2^k : then output k complete random bits..
 - ... entropy is k.
- **1** Let $2^k < m < 2^{k+1}$ biggest block.
- $ullet u = \lfloor \lg(m-2^k) \rfloor < k.$ There must be a block of size u in the decomposition of m.
- $m{0}$ two blocks in decomposition of m: sizes 2^k and 2^u .
- Largest two blocks...

- Valid extractor...

- **4** It is a block of size 2^k : then output k complete random bits..
 - ... entropy is k.
- **1** Let $2^k < m < 2^{k+1}$ biggest block.
- $ullet u = \left\lfloor \lg(m-2^k)
 ight
 floor < k.$ There must be a block of size u in the decomposition of m.
- **1** two blocks in decomposition of m: sizes 2^k and 2^u .
- Largest two blocks...
- $2^k + 2 * 2^u > m \implies 2^{u+1} + 2^k m > 0.$

- Valid extractor...
- $oldsymbol{0}$ m not a power of $oldsymbol{2}$...
- **4** It is a block of size 2^k : then output k complete random bits..
 - ... entropy is k.
- ullet Let $2^k < m < 2^{k+1}$ biggest block.
- $ullet u = \left\lfloor \lg(m-2^k)
 ight
 floor < k.$ There must be a block of size u in the decomposition of m.
- $oldsymbol{0}$ two blocks in decomposition of m: sizes 2^k and 2^u .
- Largest two blocks...
- $\mathbf{0}$ \mathbf{Y} : random variable = number of bits output by extractor.

1 By lemma, since $\frac{m-2^k}{m} < 1$:

$$\frac{m-2^k}{m} \leq \frac{m-2^k + \left(2^{u+1} + 2^k - m\right)}{m + \left(2^{u+1} + 2^k - m\right)} = \frac{2^{u+1}}{2^{u+1} + 2^k}.$$

② By induction (assumed holds for all numbers smaller than m):

$$\mathbf{E}[Y] \geq rac{2^k}{m}k + rac{m-2^k}{m}igg(\underbrace{\lfloor\lg(m-2^k)
floor}_{u}-1igg)$$

$$= rac{2^k}{m}k + rac{m-2^k}{m}(\underbrace{k-k}_{=0}+u-1)$$

$$=k+\frac{m-2^k}{m}(u-k-1)$$

1 By lemma, since $\frac{m-2^k}{m} < 1$:

$$\frac{m-2^k}{m} \leq \frac{m-2^k + \left(2^{u+1} + 2^k - m\right)}{m + \left(2^{u+1} + 2^k - m\right)} = \frac{2^{u+1}}{2^{u+1} + 2^k}.$$

② By induction (assumed holds for all numbers smaller than m):

$$egin{aligned} \mathbf{E}[Y] & \geq rac{2^k}{m}k + rac{m-2^k}{m}igg(\underbrace{\lfloor \lg(m-2^k)
floor}_{u} -1 igg) \end{aligned} \ & = rac{2^k}{m}k + rac{m-2^k}{m}\underbrace{(k-k+u-1)}_{=0} \end{aligned}$$

$$=k+\frac{m-2^k}{m}(u-k-1)$$

1 By lemma, since $\frac{m-2^k}{m} < 1$:

$$\frac{m-2^k}{m} \leq \frac{m-2^k + \left(2^{u+1} + 2^k - m\right)}{m + \left(2^{u+1} + 2^k - m\right)} = \frac{2^{u+1}}{2^{u+1} + 2^k}.$$

② By induction (assumed holds for all numbers smaller than m):

$$egin{aligned} \operatorname{E}[Y] & \geq rac{2^k}{m}k + rac{m-2^k}{m}igg(\underbrace{\lfloor \lg(m-2^k)
floor}_u - 1igg) \ & = rac{2^k}{m}k + rac{m-2^k}{m}(\underbrace{k-k}_{=0} + u - 1) \end{aligned}$$

$$=k+\frac{m-2^k}{m}(u-k-1)$$

① By lemma, since $\frac{m-2^k}{m} < 1$:

$$\frac{m-2^k}{m} \leq \frac{m-2^k+\left(2^{u+1}+2^k-m\right)}{m+\left(2^{u+1}+2^k-m\right)} = \frac{2^{u+1}}{2^{u+1}+2^k}.$$

② By induction (assumed holds for all numbers smaller than m):

$$\begin{split} \mathrm{E}[Y] & \geq \frac{2^k}{m} k + \frac{m-2^k}{m} \bigg(\underbrace{\left\lfloor \lg(m-2^k) \right\rfloor}_{u} - 1 \bigg) \\ & = \frac{2^k}{m} k + \frac{m-2^k}{m} (\underbrace{k-k}_{=0} + u - 1) \\ & = k + \frac{m-2^k}{m} (u-k-1) \end{split}$$

Sariel (UIUC)

We have:

$$egin{split} \mathbf{E}ig[Yig] & \geq k + rac{m-2^k}{m}(u-k-1) \ & \geq k + rac{2^{u+1}}{2^{u+1}+2^k}(u-k-1) \ & = k - rac{2^{u+1}}{2^{u+1}+2^k}(1+k-u), \end{split}$$

- ② If u=k-1, then $\mathrm{E}[Y] \geq k-\frac{1}{2}\cdot 2=k-1$, as required.
- \bullet If u=k-2 then $\operatorname{E}[Y] \geq k-\frac{1}{2}\cdot 3=k-1$.

New CS473 Fall 2015 29 / 30

We have:

$$egin{split} \mathrm{E}ig[Yig] &\geq k + rac{m-2^k}{m}(u-k-1) \ &\geq k + rac{2^{u+1}}{2^{u+1}+2^k}(u-k-1) \ &= k - rac{2^{u+1}}{2^{u+1}+2^k}(1+k-u), \end{split}$$

since $u-k-1 \leq 0$ as k > u.

- ② If u=k-1, then $\mathop{\mathrm{E}}[Y] \geq k \frac{1}{2} \cdot 2 = k-1$, as required.
- $extbf{3} ext{ If } u=k-2 ext{ then } ext{E}[Y] \geq k-rac{1}{3} \cdot 3 = k-1.$

We have:

$$egin{split} \mathrm{E}ig[Yig] &\geq k + rac{m-2^k}{m}(u-k-1) \ &\geq k + rac{2^{u+1}}{2^{u+1}+2^k}(u-k-1) \ &= k - rac{2^{u+1}}{2^{u+1}+2^k}(1+k-u), \end{split}$$

since $u - k - 1 \le 0$ as k > u.

- ② If u=k-1, then $\mathop{
 m E}[Y] \ge k rac{1}{2} \cdot 2 = k-1$, as required.
- $extbf{3} ext{ If } u=k-2 ext{ then } ext{E}[Y] \geq k-rac{1}{3} \cdot 3 = k-1.$

We have:

$$egin{split} \mathrm{E}ig[Yig] &\geq k + rac{m-2^k}{m}(u-k-1) \ &\geq k + rac{2^{u+1}}{2^{u+1}+2^k}(u-k-1) \ &= k - rac{2^{u+1}}{2^{u+1}+2^k}(1+k-u), \end{split}$$

since $u-k-1 \leq 0$ as k>u.

- ② If u=k-1, then $\mathop{\mathrm{E}}[Y] \geq k rac{1}{2} \cdot 2 = k-1$, as required.
- lacksquare 1 If u=k-2 then $\mathrm{E}[Y]\geq k-rac{1}{3}\cdot 3=k-1$.

Sariel (UIUC) New CS473 29 Fall 2015 29 / 30

Proof continued.....

- $ullet {f E}[Y] \geq k rac{2^{u+1}}{2^{u+1}+2^k}(1+k-u).$ And $u-k-1 \leq 0$ as k>u.
- $oldsymbol{2}$ If u < k-2 then

$$\begin{split} \mathrm{E}[Y] & \geq k - \frac{2^{u+1}}{2^k} (1+k-u) \\ & = k - \frac{k-u+1}{2^{k-u-1}} \\ & = k - \frac{2+(k-u-1)}{2^{k-u-1}} \\ & \geq k-1, \end{split}$$

since $(2+i)/2^i \le 1$ for $i \ge 2$.

Proof continued.....

- $egin{array}{l} \bullet & \mathrm{E}[Y] \geq k rac{2^{u+1}}{2^{u+1}+2^k} (1+k-u). \ & \mathrm{And} \; u-k-1 \leq 0 \; \mathrm{as} \; k > u. \end{array}$
- $oldsymbol{0}$ If u < k-2 then

$$egin{aligned} \mathrm{E}[Y] & \geq k - rac{2^{u+1}}{2^k} (1+k-u) \ & = k - rac{k-u+1}{2^{k-u-1}} \ & = k - rac{2+(k-u-1)}{2^{k-u-1}} \ & \geq k-1, \end{aligned}$$

since $(2+i)/2^i \le 1$ for $i \ge 2$.

Sariel (UIUC) New CS473 31 Fall 2015 31 / 30

Sariel (UIUC) New CS473 32 Fall 2015 32 / 30

Sariel (UIUC) New CS473 33 Fall 2015 33 / 30

Sariel (UIUC) New CS473 34 Fall 2015 34 / 30