# **Entropy, Randomness, and Information**

Lecture 23 November 13, 2014

### 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,  $\mathbb{H}(X) = \mathrm{E} \left[ \lg \frac{1}{\Pr[X]} \right]$ .

#### 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?

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

#### 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?

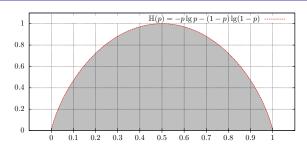
$$\mathbb{H}(X) = -\sum_x \Prig[X = xig] \lg \Prig[X = xig]$$

#### 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  $\mathbb{H}(0) = \mathbb{H}(1) = 0$ .

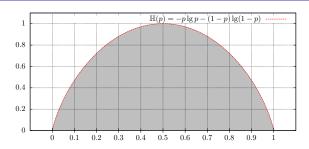
Q: How many truly random bits are there when given the result of flipping a single coin with probability p for heads?

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



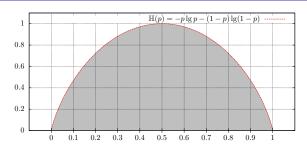
- $lackbox{1}{} \mathbb{H}(p)$  is a concave symmetric around 1/2 on the interval [0,1].
- $\bigcirc$  maximum at 1/2.
- **3**  $\mathbb{H}(3/4) \approx 0.8113$  and  $\mathbb{H}(7/8) \approx 0.5436$ .
- $\implies$  coin that has 3/4 probably to be heads have higher amount of "randomness" in it than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

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



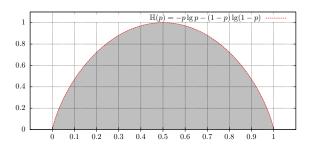
- $lackbox{1}{} \mathbb{H}(p)$  is a concave symmetric around 1/2 on the interval [0,1].
- $\bigcirc$  maximum at 1/2.
- **3**  $\mathbb{H}(3/4) \approx 0.8113$  and  $\mathbb{H}(7/8) \approx 0.5436$ .
- $\Longrightarrow$  coin that has 3/4 probably to be heads have higher amount of "randomness" in it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

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



- $lackbox{1}{} \mathbb{H}(p)$  is a concave symmetric around 1/2 on the interval [0,1].
- $oldsymbol{0}$  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.

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



- $lackbox{1}{} \mathbb{H}(p)$  is a concave symmetric around 1/2 on the interval [0,1].
- $oldsymbol{0}$  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.

- ②  $\mathbb{H}'(p) = -\lg p + \lg(1-p) = \lg \frac{1-p}{p}$
- $\blacksquare \implies \mathbb{H}''(p) \leq 0$ , for all  $p \in (0,1)$ , and the  $\mathbb{H}(\cdot)$  is concave.
- balanced coin has the largest amount of randomness in it

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

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

- **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

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

- **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.

...out of bad random bits...

- 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!
- $\bigcirc$  p: probability for head.
- 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?

- 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!
- We need fair bit coins!
- ${ ilde {f 0}}$  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?

...out of bad random bits...

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

Sariel (UIUC) CS573 Fall 2014

- $b_1, \ldots, b_n$ : result of n coin flips...
- From a faulty coin!
- p: probability for head.
- We need fair bit coins!
- **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!
- **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  $pprox n\mathbb{H}(p)$ .

- **1** entropy of X is  $\mathbb{H}(X) = -\sum_x \Pr[X = x] \lg \Pr[X = x]$ .

 $i=1,\ldots,n$ , has entropy  $\mathbb{H}(X)=-\sum_{i=1}^n\frac{1}{n}\lg\frac{1}{n}=\lg n$ .

CS573 Fall 2014 11 / 30

- **①** *entropy* of X is  $\mathbb{H}(X) = -\sum_x \Prig[X = xig] \lg \Prig[X = xig]$  .
- 2 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.

- **①** *entropy* 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.

- **1** entropy 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.

- **①** *entropy* of X is  $\mathbb{H}(X) = -\sum_x \Prig[X = xig] \lg \Prig[X = xig].$
- 2 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

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

#### Proof continued

$$\begin{split} \mathbb{H}(Z) &= \sum_{x} \sum_{y} \Pr[X=x] \Pr[Y=y] \lg \frac{1}{\Pr[X=x]} \\ &+ \sum_{y} \sum_{x} \Pr[X=x] \Pr[Y=y] \lg \frac{1}{\Pr[Y=y]} \\ &= \sum_{x} \Pr[X=x] \lg \frac{1}{\Pr[X=x]} \\ &+ \sum_{y} \Pr[Y=y] \lg \frac{1}{\Pr[Y=y]} \\ &= \mathbb{H}(X) + \mathbb{H}(Y) \,. \end{split}$$

Sariel (UIUC) CS573 14 Fall 2014 14 / 30

### Bounding the binomial coefficient using entropy

#### Lemma

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

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

#### **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

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

#### Proof

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

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

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

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

#### 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

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

#### 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

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

- ②  $\sum_{i=0}^{n} {n \choose i} q^i (1-q)^{n-i} = \sum_{i=0}^{n} \mu(i)$ .
- ullet Claim:  $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$  largest term in  $\sum_{k=0}^n \mu(k)=1.$
- $oldsymbol{\Phi} \Delta_k = \mu(k) \mu(k+1) = inom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q}\Big),$
- lacksquare sign of  $\Delta_k = \mathsf{size}$  of last term..
- $egin{align} \mathbf{0} & ext{sign}(oldsymbol{\Delta}_k) = ext{sign}\Big(1-rac{(n-k)\,q}{(k+1)(1-q)}\Big) \ & = ext{sign}\Big(rac{(k+1)(1-q)-(n-k)\,q}{(k+1)(1-q)}\Big) \end{aligned}$

② 
$$\sum_{i=0}^{n} {n \choose i} q^{i} (1-q)^{n-i} = \sum_{i=0}^{n} \mu(i)$$
.

- ullet Claim:  $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$  largest term in  $\sum_{k=0}^n \mu(k)=1.$
- $oldsymbol{\Phi}_k = \mu(k) \mu(k+1) = \binom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q}\Big),$
- ullet sign of  $\Delta_k=$  size of last term...
- $\begin{array}{l} \mathbf{6} \ \operatorname{sign}(\Delta_k) = \operatorname{sign}\!\left(1 \frac{(n-k)q}{(k+1)(1-q)}\right) \\ = \operatorname{sign}\!\left(\frac{(k+1)(1-q) (n-k)q}{(k+1)(1-q)}\right). \end{array}$

$$\bullet \mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$$

② 
$$\sum_{i=0}^{n} {n \choose i} q^{i} (1-q)^{n-i} = \sum_{i=0}^{n} \mu(i)$$
.

- ullet Claim:  $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$  largest term in  $\sum_{k=0}^n \mu(k)=1.$
- $oldsymbol{\Phi} \Delta_k = \mu(k) \mu(k+1) = inom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q}\Big),$
- sign of  $\Delta_k =$  size of last term...
- $egin{align} \mathbf{0} & ext{sign}(oldsymbol{\Delta}_k) = ext{sign}\Big(1-rac{(n-k)\,q}{(k+1)(1-q)}\Big) \ & = ext{sign}\Big(rac{(k+1)(1-q)-(n-k)\,q}{(k+1)(1-q)}\Big) \end{aligned}$

$$\mathbf{0} \ \mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$$

② 
$$\sum_{i=0}^{n} {n \choose i} q^{i} (1-q)^{n-i} = \sum_{i=0}^{n} \mu(i)$$
.

- ullet Claim:  $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$  largest term in  $\sum_{k=0}^n \mu(k)=1.$
- $oldsymbol{\Delta}_k = \mu(k) \mu(k+1) = inom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q}\Big),$
- $\begin{aligned} \mathbf{0} \ \ & \mathrm{sign}(\boldsymbol{\Delta}_k) = \mathrm{sign}\Big(1 \frac{(n-k)q}{(k+1)(1-q)}\Big) \\ & = \mathrm{sign}\Big(\frac{(k+1)(1-q) (n-k)q}{(k+1)(1-q)}\Big) \end{aligned}$

$$\mathbf{0} \ \mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$$

② 
$$\sum_{i=0}^{n} {n \choose i} q^{i} (1-q)^{n-i} = \sum_{i=0}^{n} \mu(i)$$
.

- ullet Claim:  $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$  largest term in  $\sum_{k=0}^n \mu(k)=1.$
- $oldsymbol{\Delta}_k = \mu(k) \mu(k+1) = inom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q}\Big),$
- $\bullet$  sign of  $\Delta_k =$  size of last term...
- $egin{align} \mathbf{0} & ext{sign}(oldsymbol{\Delta}_k) = ext{sign}\Big(1-rac{(n-k)\,q}{(k+1)(1-q)}\Big) \ & = ext{sign}\Big(rac{(k+1)(1-q)-(n-k)\,q}{(k+1)(1-q)}\Big) \end{aligned}$

$$\mathbf{0} \ \mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$$

② 
$$\sum_{i=0}^{n} {n \choose i} q^{i} (1-q)^{n-i} = \sum_{i=0}^{n} \mu(i)$$
.

- ullet Claim:  $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$  largest term in  $\sum_{k=0}^n \mu(k)=1.$
- $oldsymbol{\Delta}_k = \mu(k) \mu(k+1) = inom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q}\Big),$
- lacktriangledown sign of  $\Delta_k = \text{size}$  of last term...
- $egin{align} \mathbf{0} & \operatorname{sign}(\Delta_k) = \operatorname{sign}ig(1 rac{(n-k)q}{(k+1)(1-q)}ig) \ & = \operatorname{sign}ig(rac{(k+1)(1-q) (n-k)q}{(k+1)(1-q)}ig) \end{aligned}$

$$\mathbf{0} \ \mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$$

② 
$$\sum_{i=0}^{n} {n \choose i} q^{i} (1-q)^{n-i} = \sum_{i=0}^{n} \mu(i)$$
.

- ullet Claim:  $\mu(nq)=inom{n}{nq}q^{nq}(1-q)^{n-nq}$  largest term in  $\sum_{k=0}^n \mu(k)=1.$
- $oldsymbol{\Delta}_k = \mu(k) \mu(k+1) = inom{n}{k} q^k (1-q)^{n-k} \Big(1 rac{n-k}{k+1} rac{q}{1-q}\Big),$
- lacktriangledown sign of  $lacktriangledown_k$  = size of last term...
- $\begin{array}{l} \bullet \ \operatorname{sign}(\Delta_k) = \operatorname{sign}\left(1 \frac{(n-k)q}{(k+1)(1-q)}\right) \\ = \operatorname{sign}\left(\frac{(k+1)(1-q) (n-k)q}{(k+1)(1-q)}\right). \end{array}$

- ②  $\Longrightarrow \Delta_k \geq 0$  when  $k \geq nq+q-1$   $\Delta_k < 0$  otherwise.

- $\Longrightarrow \ \mu(nq) \$ is the largest term in  $\sum_{k=0}^n \mu(k) = 1.$
- ullet  $\mu(nq)$  larger than the average in sum.

- $egin{array}{ll} igotimes \Delta_k \geq 0 & ext{when } k \geq nq+q-1 \ \Delta_k < 0 & ext{otherwise}. \end{array}$

- 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.

- $egin{aligned} igotimes \Delta_k & \geq 0 ext{ when } k \geq nq+q-1 \ \Delta_k & < 0 ext{ otherwise}. \end{aligned}$

- 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.

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

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

- $egin{aligned} igotimes \Delta_k & \geq 0 ext{ when } k \geq nq+q-1 \ \Delta_k & < 0 ext{ otherwise}. \end{aligned}$
- $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- 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.

- $egin{aligned} igotimes \Delta_k & \geq 0 ext{ when } k \geq nq+q-1 \ \Delta_k & < 0 ext{ otherwise}. \end{aligned}$
- **3**  $\mu(k) = \binom{n}{k} q^k (1-q)^{n-k}$
- 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.
- $lacksquare \implies inom{n}{nq} \geq rac{1}{n+1} q^{-nq} (1-q)^{-(n-nq)} = rac{1}{n+1} 2^{n\mathbb{H}(q)}.$

#### Generalization...

#### Corollary

We have:

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

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

Proof is straightforward but tedious.

- 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  $\binom{n}{nq} pprox \mathbf{2}^{n\mathbb{H}(q)}$  possible sequences .
  - (IV) ...of similar probability.
  - $(V) \implies \mathbb{H}(Y) = n\mathbb{H}(q) \approx \lg \binom{n}{nq}.$

- 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  $\binom{n}{nq} pprox \mathbf{2}^{n\mathbb{H}(q)}$  possible sequences .
  - (IV) ...of similar probability.
  - $(V) \implies \mathbb{H}(Y) = n\mathbb{H}(q) \approx \lg \binom{n}{nq}.$

- 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  $\binom{n}{nq} pprox \mathbf{2}^{n\mathbb{H}(q)}$  possible sequences .
  - (IV) ...of similar probability.
  - $(\mathsf{V}) \implies \mathbb{H}\left(Y\right) = n\mathbb{H}\left(q\right) pprox \lg \binom{n}{nq}.$

- 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  $\binom{n}{nq} pprox \mathbf{2}^{n\mathbb{H}(q)}$  possible sequences .
  - (IV) ...of similar probability.
  - $(V) \implies \mathbb{H}(Y) = n\mathbb{H}(q) \approx \lg \binom{n}{nq}.$

- 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  $\binom{n}{nq} pprox \mathbf{2}^{n\mathbb{H}(q)}$  possible sequences .
  - (IV) ...of similar probability.
  - $(\mathsf{V}) \implies \mathbb{H}\left(Y
    ight) = n\mathbb{H}\left(q
    ight) pprox \lginom{n}{nq}.$

#### 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 **Ext** takes as input the value of a random variable X and outputs a sequence of bits y, such that  $\Pr\left[\mathsf{Ext}(X) = y \ \middle| \ |y| = k\right] = \frac{1}{2^k}$ , 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  $0 \le x \le 7$

  - ① Idea... Output binary representation of x-8 as a two bit number.
- 5 A valid extractor...

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

- **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  $0 \le x \le 7$

  - lacktriangled Idea... Output binary representation of x-8 as a two bit number.
- 6 A valid extractor...

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

- **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  $0 \le x \le 7$ .

  - ① Idea... Output binary representation of x-8 as a two bit number.
- A valid extractor...

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

- **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  $0 \le x \le 7$ .

  - ① 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  $0 \le x \le 7$ .

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

- **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.
- $egin{aligned} oldsymbol{\Im} & ext{A valid extractor...} \ & ext{Pr}ig[\mathsf{Ext}(X) = 00 \ ig| \ |\mathsf{Ext}(X)| = 2 ig] = rac{1}{4}, \end{aligned}$

- **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.
- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- **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  $0 \le x \le 7$ .
  - 3 If x is between 8 and 11?
  - **1** Idea... Output binary representation of x-8 as a two bit number.
- A valid extractor...

$$\Prigl[\mathsf{Ext}(X) = 00 \ igr| \, |\mathsf{Ext}(X)| = 2 igr] = rac{1}{4},$$

#### 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  $\{0,\ldots,m-1\}$ .
- ② Then there is an extraction function for X:
  - outputs on average at least

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

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

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

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

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

independent and unbiased bits.

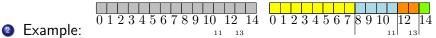
- $oldsymbol{m}$ : A sum of unique powers of  $oldsymbol{2}$ , namely  $m=\sum_i a_i 2^i$ , where  $a_i \in \{0,1\}$ .
- Example:
- ① decomposed  $\{0, \ldots, m-1\}$  into disjoint union of blocks sizes are powers of 2.
- ullet If x is in block  $2^k$ , output its relative location in the block in binary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



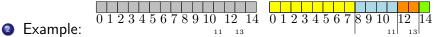
- ② Example:
- ① decomposed  $\{0, \ldots, m-1\}$  into disjoint union of blocks sizes are powers of 2.
- If x is in block  $2^k$ , output its relative location in the block in binary representation.
- Example: x = 10: then falls into block 2<sup>2</sup>... x relative location is 2. Output 2 written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

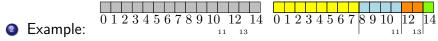


- decomposed  $\{0,\ldots,m-1\}$  into disjoint union of blocks sizes are powers of .
- ullet If x is in block  $2^k$ , output its relative location in the block in binary representation.



- 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  $a_i \in \{0,1\}$ .

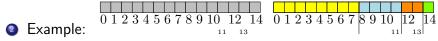


- ullet decomposed  $\{0,\ldots,m-1\}$  into disjoint union of blocks sizes are powers of 2.
- ullet If x is in block  $2^k$ , output its relative location in the block in binary representation.



- 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 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".

- 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  $\mathbf{2}^k$ : then output k complete random bits... entropy is k.
- **6** 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  $2^u$ .
- Largest two blocks...
- Y: random variable = number of bits output by extractor.

- Valid extractor...
- @ 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  $\mathbf{2}^k$ : then output k complete random bits... entropy is k.
- ullet Let  $2^k < m < 2^{k+1}$  biggest block.
- $\bigcirc$  two blocks in decomposition of m: sizes  $2^k$  and  $2^u$ .
- Largest two blocks...
- $\mathbf{O}$  Y: random variable = number of bits output by extractor.

- Valid extractor...
- ② Theorem holds if m is a power of two. Only one block.
- ullet X falls in block of size  $\mathbf{2}^k$ : then output k complete random bits.. ... entropy is k.
- ullet Let  $2^k < m < 2^{k+1}$  biggest block.
- $\bigcirc$  two blocks in decomposition of m: sizes  $2^k$  and  $2^u$ .
- Largest two blocks...
- $\mathbf{O}$  Y: random variable = number of bits output by extractor.

- Valid extractor...
- ② Theorem holds if m is a power of two. Only one block.
- $oldsymbol{0}$  m not a power of  $oldsymbol{2}$ ...
- igoplus X falls in block of size  $2^k$ : then output k complete random bits... entropy is k.
- **6** Let  $2^k < m < 2^{k+1}$  biggest block.
- $\bigcirc$  two blocks in decomposition of m: sizes  $2^k$  and  $2^u$ .
- Largest two blocks...

- Valid extractor...
- ② Theorem holds if m is a power of two. Only one block.
- **4** 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.
- 0  $u = \lfloor \lg(m-2^k) \rfloor < k$ . 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.
- **4** X falls in block of size  $2^k$ : then output k complete random bits... entropy is k.
- lacksquare 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.
- $\bigcirc$  two blocks in decomposition of m: sizes  $2^k$  and  $2^u$ .
- Largest two blocks...

- Valid extractor...
- ② Theorem holds if m is a power of two. Only one block.
- **4** X falls in block of size  $2^k$ : then output k complete random bits... entropy is k.
- lacksquare Let  $2^k < m < 2^{k+1}$  biggest block.
- ① two blocks in decomposition of m: sizes  $2^k$  and  $2^u$ .
- 8 Largest two blocks...
- $oldsymbol{0}$  Y: random variable = number of bits output by extractor.

- Valid extractor...
- Theorem holds if m is a power of two. Only one block.
- **4** If alls in block of size  $2^k$ : then output k complete random bits... entropy is k.
- lacksquare Let  $2^k < m < 2^{k+1}$  biggest block.
- $oldsymbol{0}$  two blocks in decomposition of m: sizes  $2^k$  and  $2^u$ .
- Largest two blocks...

- Valid extractor...
- ② Theorem holds if m is a power of two. Only one block.
- **4** X falls in block of size  $2^k$ : then output k complete random bits... entropy is k.
- **5** Let  $2^k < m < 2^{k+1}$  biggest block.
- $oldsymbol{0}$  two blocks in decomposition of m: sizes  $2^k$  and  $2^u$ .
- Largest two blocks...

- Valid extractor...
- Theorem holds if m is a power of two. Only one block.
- **4** If alls in block of size  $2^k$ : then output k complete random bits... entropy is k.
- lacksquare Let  $2^k < m < 2^{k+1}$  biggest block.
- lacksquare two blocks in decomposition of m: sizes  $2^k$  and  $2^u$ .
- Largest two blocks...
- $\bigcirc Y$ : random variable = number of bits output by extractor.

- Valid extractor...
- Theorem holds if m is a power of two. Only one block.
- $oldsymbol{0}$  m not a power of  $oldsymbol{2}$ ...
- **4** If alls in block of size  $2^k$ : then output k complete random bits... entropy is k.
- lacksquare Let  $2^k < m < 2^{k+1}$  biggest block.
- lacksquare two blocks in decomposition of m: sizes  $2^k$  and  $2^u$ .
- Largest two blocks...
- $oldsymbol{0}$  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):

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

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

**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):

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

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

**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):

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

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

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

$$rac{m-2^k}{m} \leq rac{m-2^k+ig(2^{u+1}+2^k-mig)}{+(2^{u+1}+2^k-m)} = rac{2^{u+1}}{2^{u+1}+2^k}.$$

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

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

We have:

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

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

CS573 Fall 2014 29 / 30

We have:

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

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:

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

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

- **1** If u = k 2 then  $\mathbf{E}[Y] > k \frac{1}{2} \cdot 3 = k 1$ .

Sariel (UIUC) CS573 29 Fall 2014 29 / 30

We have:

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

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  $\mathrm{E}[Y]\geq k-rac{1}{3}\cdot 3=k-1$ .

### Proof continued.....

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

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

since  $(2+i)/2^i \leq 1$  for  $i \geq 2$ .

### Proof continued.....

- $egin{array}{l} \bullet & \mathrm{E}[\,Y] \geq k rac{2^{u+1}}{2^{u+1}+2^k} (1+k-u). \ & \mathsf{And} \,\, u-k-1 \leq 0 \,\, \mathsf{as} \,\, k > u. \end{array}$
- 2 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  $(2+i)/2^i \leq 1$  for  $i \geq 2$ .

Sariel (UIUC) CS573 31 Fall 2014 31 / 30

Sariel (UIUC) CS573 33 Fall 2014 33 / 30

Sariel (UIUC) CS573 34 Fall 2014 34 / 30