Lecture 22

General Question:
What is the average codeword length of unary codes (without using Golomb codes)?

n	$G(n) = p^n(1-p)$	Unary Code	Codeword Length
0	$1-p$	0	1
1	$p(1-p)$	10	2
2	$p^2(1-p)$	110	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	$p^k(1-p)$	$\underbrace{111\cdots1}_{k\ 1\text{'s}}0$	$(k+1)$
$\vdots$			

$$ACL = \sum_{k=0}^{\infty} \left[p^k(1-p) \cdot (k+1) \right]$$

$$= (1-p) \sum_{k=0}^{\infty} \left[p^k \cdot (k+1) \right] = \frac{1-p}{p} \sum_{k=0}^{\infty} \left[p^{(k+1)} (k+1) \right]$$

Let $s = k+1 \Rightarrow k = s-1$

$$= \frac{(1-p)}{p} \sum_{s=1}^{\infty} \left(p^s \cdot s \right)$$

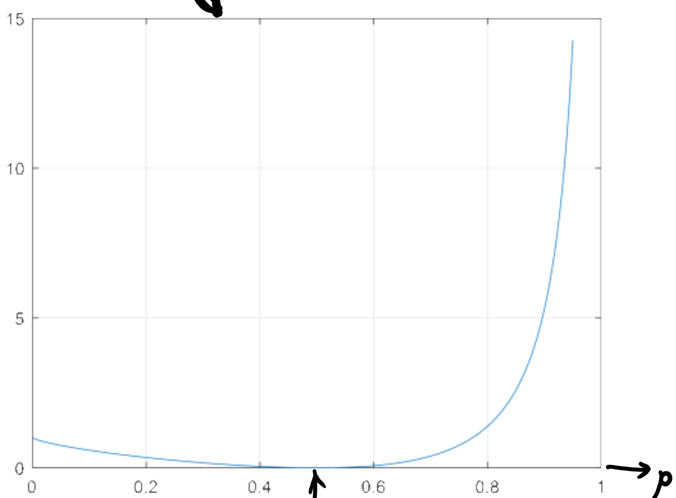
$$= \frac{(1-p)}{p} \cdot \frac{p}{(1-p)^2}$$

$$\boxed{ACL = \frac{1}{1-p}}$$

Compared with:

$$\boxed{H(X) = \frac{-p\log_2 p - (1-p)\log_2(1-p)}{1-p}} = \frac{H(p, 1-p)}{1-p}$$

Redundancy = ACL - H(X) = $\dfrac{1 - H(p, 1-p)}{1-p}$



$p = \frac{1}{2}$, Redundancy = 0.

```
>> p = 0.0001: 0.0001: 0.95;
>> H = (-p.*log2(p) - (1 - p).*log2(1-p))./(1 - p);
>> ACL = 1./(1 - p);
>> Redun = ACL - H;
>> plot(p, Redun); grid
```

Golomb code: $\boxed{p^m = \frac{1}{2}}$.

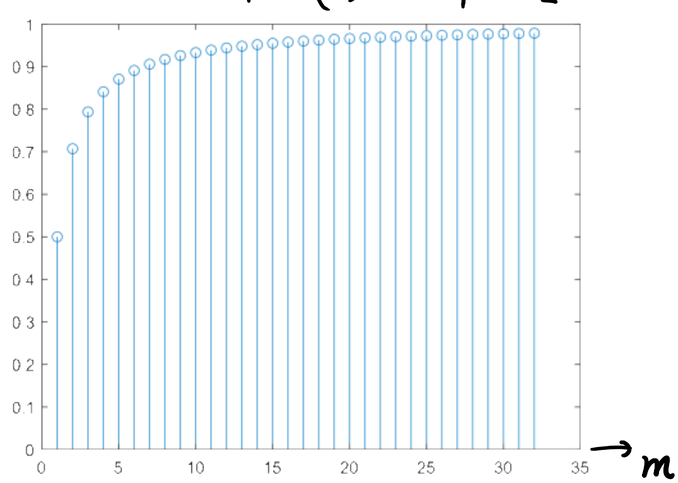
$\Downarrow$

$p = \left(\frac{1}{2}\right)^{\frac{1}{m}}$

$m \uparrow \rightarrow p \uparrow$ (>0.5)

In this special case, the Golomb code becomes the unary code.

$p = \left(\frac{1}{2}\right)^{\frac{1}{m}} \Rightarrow p^m = \frac{1}{2}$.




```
>> m = 1: 32;
>> p = (1/2).^(1./m);
>> figure; stem(m, p);
```

Unary code performs worse and worse when m increases.
Next, analyze the Golomb coding efficiency.

Code the non-negative integers n = mq + r, where m is a coding parameter (positive integer).

Split the integer n into two parts:

(1) Code **q** with unary code. Here q is the quotient of (n/m).

Unary code: q 1's, followed by one '0'. Codeword length of this unary code: (q + 1) bits

(2) Code **r** using binary code. Here r is the remainder of (n/m). Binary code has $(\log_2 m)$ bits

Example: $m = 16$

$m = 16$			
n	Codeword	n	Codeword
0	00000	24	101000
1	00001	25	101001
2	00010	26	101010
3	00011	27	101011
4	00100	28	101100
5	00101	29	101101
6	00110	30	101110
7	00111	31	101111
8	01000		
9	01001	32	1100000
10	01010	33	1100001
11	01011	34	1100010
12	01100	35	1100011
13	01101	36	1100100
14	01110	37	1100101
15	01111	38	1100110
		39	1100111
16	100000	40	1101000
17	100001	41	1101001
18	100010	42	1101010
19	100011	43	1101011
20	100100	44	1101100
21	100101	45	1101101
22	100110	46	1101110
23	100111	47	1101111

$n = 0, 1, \cdots, 15$

$\frac{n}{m} = \frac{n}{16} \Rightarrow \begin{cases} q = 0 : & \text{1-bit unary code.} \\ r = n : & \text{4-bit binary code.} \end{cases}$

If $n = j \cdot 16 + 1$, where $j = 0, 1, \cdots,$

$= 1, 17, 33, \cdots$

$\frac{n}{m} \Rightarrow r = 1$ same remainder

$$\text{Prob}[r = 1] = \sum_{j=0}^{\infty} p^{j \cdot 16 + 1}(1-p)$$

$$= (1-p)p \sum_{j=0}^{\infty} p^{16j}$$

In general, if $n = jm + r$, where $0 \leq r < m$, then all these n values lead to the same remainder r, since $n \bmod m = r$.

Given $G(n) = p^n(1-p)$,

$$\text{Prob}[\text{Remainder} = r] = \sum_{j=0}^{\infty} p^{jm+r}(1-p) = p^r(1-p)\sum_{j=0}^{\infty}(p^m)^j = \frac{p^r(1-p)}{1-p^m}$$

In summary,

$$\boxed{\text{Prob}[\text{ Remainder} = r] = \frac{p^r(1-p)}{1-p^m}}$$

For example, $m = 16$, $p^m = p^{16} = \frac{1}{2} \Rightarrow p = \left(\frac{1}{2}\right)^{\frac{1}{16}}$

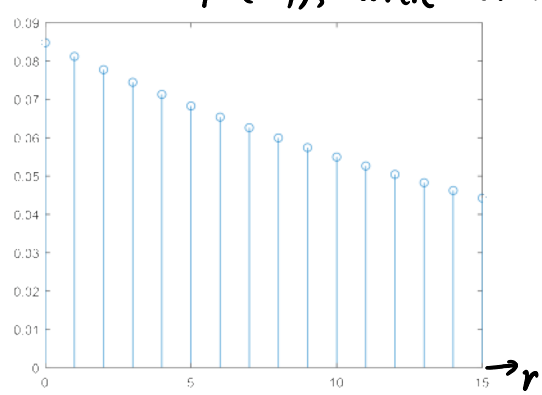
$$\text{Prob}[r] = 2p^r(1-p)$$

```
>> p = (1/2).^(1./16)
p =
   0.9576
```

```
>> p = (1/2).^(1./16)
p =
   0.9576

>> P = 2*(p.^r)*(1-p);
>> P16 = P(1:16);
>> sum(P16)
ans =
   1.0000
```

$2p^r(1-p)$, where $0 \le r < 16$, $p = 0.9576$ $(m=16)$




```
>> sum(-P16.
*log2(P16))
ans =
   3.9716
```
(Entropy)

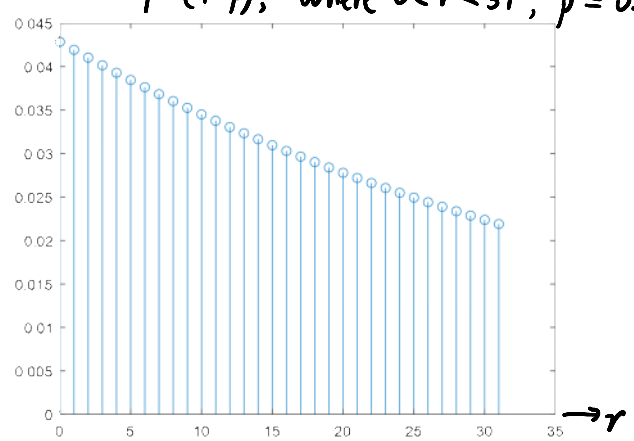
$(\log_2 m) = \log_2 16 = 4$ bits

Redundancy ≈ 0.03 bit

```
>> r = 0: 31;
>> P = 2*(p.^r)*(1-p);
>> p = (1/2).^(1./32)
p =
   0.9786

>> r = 0: 31;
>> P = 2*(p.^r)*(1-p);
>> sum(P)
ans =
   1.0000
>> sum(-P.*log2(P))
ans =
   4.9715
```
$\leftarrow$ $H(r)$ Entropy

$2p^r(1-p)$, where $0 \le r < 31$, $p = 0.9786$ $(m = 32)$



Compared to $(\log_2 m) = (\log_2 32) = 5$ bit

Redundancy ≈ 0.03 bit/ remainder value