




WIKIPEDIA
The Free Encyclopedia

[Main page](#)
[Contents](#)
[Featured content](#)
[Current events](#)
[Random article](#)
[Donate to Wikipedia](#)
[Wikipedia store](#)

Interaction
[Help](#)
[About Wikipedia](#)
[Community portal](#)
[Recent changes](#)
[Contact page](#)

Tools
[What links here](#)
[Related changes](#)
[Upload file](#)
[Special pages](#)
[Permanent link](#)
[Page information](#)
[Wikidata item](#)
[Cite this page](#)


Print/export
[Create a book](#)
[Download as PDF](#)
[Printable version](#)

Languages
[Српски / srpski](#)
 [Edit links](#)

[Create account](#) [Log in](#)

Article [Talk](#)

[Read](#) [Edit](#) [View history](#)



Shannon–Fano–Elias coding

From Wikipedia, the free encyclopedia

In [information theory](#), **Shannon–Fano–Elias coding** is a precursor to [arithmetic coding](#), in which probabilities are used to determine codewords.^[1]

Contents [hide]

- 1 [Algorithm description](#)
 - 1.1 [Example](#)
- 2 [Algorithm analysis](#)
 - 2.1 [Prefix code](#)
 - 2.2 [Code length](#)
- 3 [References](#)

Algorithm description [edit]

Given a [discrete random variable](#) *X* of [ordered](#) values to be encoded, let *p*(*x*) be the probability for any *x* in *X*.

Define a function

$$\bar{F}(x) = \sum_{x_i < x} p(x_i) + \frac{1}{2}p(x)$$

Algorithm:

For each *x* in *X*,

Let *Z* be the binary expansion of $\bar{F}(x)$.

Choose the length of the encoding of *x*, *L*(*x*), to be the integer $\left\lceil \log_2 \frac{1}{p(x)} \right\rceil + 1$

Choose the encoding of *x*, *code*(*x*), be the first *L*(*x*) [most significant bits](#) after the decimal point of *Z*.

Example [edit]

Let *X* = {*A*, *B*, *C*, *D*}, with probabilities *p* = {1/3, 1/4, 1/6, 1/4}.

For *A*

$$\bar{F}(A) = \frac{1}{2}p(A) = \frac{1}{2} \cdot \frac{1}{3} = 0.1666...$$

In binary, *Z*(*A*) = 0.0010101010...

$$L(A) = \left\lceil \log_2 \frac{1}{\frac{1}{3}} \right\rceil + 1 = 3$$

code(*A*) is 001

For *B*

$$\bar{F}(B) = p(A) + \frac{1}{2}p(B) = \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{4} = 0.4583333...$$

In binary, *Z*(*B*) = 0.01110101010101...

$$L(B) = \left\lceil \log_2 \frac{1}{\frac{1}{4}} \right\rceil + 1 = 3$$

code(*B*) is 011

For *C*

$$\bar{F}(C) = p(A) + p(B) + \frac{1}{2}p(C) = \frac{1}{3} + \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{6} = 0.666666...$$

In binary, *Z*(*C*) = 0.101010101010...

$$L(C) = \left\lceil \log_2 \frac{1}{\frac{1}{6}} \right\rceil + 1 = 4$$

code(*C*) is 1010

For D

$$\bar{F}(D) = p(A) + p(B) + p(C) + \frac{1}{2}p(D) = \frac{1}{3} + \frac{1}{4} + \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{4} = 0.875$$

In binary, $Z(D) = 0.111$

$$L(D) = \left\lceil \log_2 \frac{1}{\frac{1}{4}} \right\rceil + 1 = 3$$

code(D) is 111

Algorithm analysis [\[edit\]](#)

Prefix code [\[edit\]](#)

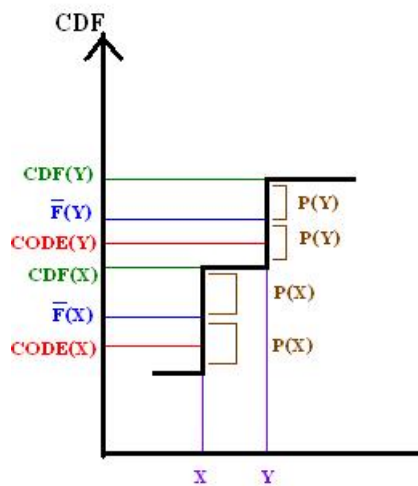
Shannon–Fano–Elias coding produces a binary [prefix code](#), allowing for direct decoding.

Let $bcode(x)$ be the rational number formed by adding a decimal point before a binary code. For example, if $code(C)=1010$ then $bcode(C) = 0.1010$. For all x , if no y exists such that

$$bcode(x) \leq bcode(y) < bcode(x) + 2^{-L(x)}$$

then all the codes form a prefix code.

By comparing F to the [CDF](#) of X , this property may be demonstrated graphically for Shannon–Fano–Elias coding.



By definition of L it follows that

$$2^{-L(x)} \leq \frac{1}{2}p(x)$$

And because the bits after $L(y)$ are truncated from $F(y)$ to form $code(y)$, it follows that

$$\bar{F}(y) - bcode(y) \leq 2^{-L(y)}$$

thus $bcode(y)$ must be no less than $CDF(x)$.

So the above graph demonstrates that the $bcode(y) - bcode(x) > p(x) \geq 2^{-L(x)}$, therefore the prefix property holds.

Code length [\[edit\]](#)

The average code length is $LC(X) = \sum_{x \in X} p(x)L(x) = \sum_{x \in X} p(x) \left(\left\lceil \log_2 \frac{1}{p(x)} \right\rceil + 1 \right)$.



Thus for $H(X)$, the [Entropy](#) of the random variable X ,

$$H(X) + 1 \leq LC(X) < H(X) + 2$$

Shannon Fano Elias codes from 1 to 2 extra bits per symbol from X than entropy, so the code is not used in practice.

References [\[edit\]](#)

- [▲] T. M. Cover and Joy A. Thomas (2006). *Elements of information theory* (2nd ed.). John Wiley and Sons. pp. 127–128. ISBN 978-0-471-24195-9.

v · t · e		Data compression methods	[hide]
Lossless	Entropy type	Unary · Arithmetic · Golomb · Huffman (Adaptive · Canonical · Modified) · Range · Shannon · Shannon–Fano · Shannon–Fano–Elias · Tunstall · Universal (Exp-Golomb · Fibonacci · Gamma · Levenshtein)	
	Dictionary type	Byte pair encoding · DEFLATE · Lempel–Ziv (LZ77 / LZ78 (LZ1 / LZ2) · LZJB · LZMA · LZO · LZRW · LZS · LZSS · LZW · LZWL · LZX · LZ4 · Statistical)	
	Other types	BWT · CTW · Delta · DMC · MTF · PAQ · PPM · RLE	
Audio	Concepts	Bit rate (average (ABR) · constant (CBR) · variable (VBR)) · Companding · Convolution · Dynamic range · Latency · Nyquist–Shannon theorem · Sampling · Sound quality · Speech coding · Sub-band coding	
	Codec parts	A-law · μ-law · ACELP · ADPCM · CELP · DPCM · Fourier transform · LPC (LAR · LSP) · MDCT · Psychoacoustic model · WLPc	
Image	Concepts	Chroma subsampling · Coding tree unit · Color space · Compression artifact · Image resolution · Macroblock · Pixel · PSNR · Quantization · Standard test image	
	Methods	Chain code · DCT · EZW · Fractal · KLT · LP · RLE · SPIHT · Wavelet	
Video	Concepts	Bit rate (average (ABR) · constant (CBR) · variable (VBR)) · Display resolution · Frame · Frame rate · Frame types · Interlace · Video characteristics · Video quality	
	Codec parts	Lapped transform · DCT · Deblocking filter · Motion compensation	
Theory	Entropy · Kolmogorov complexity · Lossy · Quantization · Rate–distortion · Redundancy · Timeline of information theory		
<div><div> Compression formats</div><div> Compression software (codecs)</div></div>			

Categories: Lossless compression algorithms

This page was last modified on 2 June 2014, at 20:21.

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. By using this site, you agree to the [Terms of Use](#) and [Privacy Policy](#). Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc.](#), a non-profit organization.

[Privacy policy](#) [About Wikipedia](#) [Disclaimers](#) [Contact Wikipedia](#) [Developers](#) [Mobile view](#)

