Information Theory
CSE 468 Fall 2025

jedimaestro@asu.edu

Don’t do this...

```
int i = 10000000 + new Random().nextInt(89999999);
int j = 10000000 + new Random().nextInt(89999999);
return (String.valueOf(i) + String.valueOf(j)).getBytes();
```

Figure 1: Decompiled Java method generating an AES session key in version 6.3.0.1920.

```
Random random = new Random(System.currentTimeMillis());
byte[] bArr = new byte[8];
byte[] bArr2 = new byte[8];
random.nextBytes(bArr);
random.nextBytes(bArr2);
return new SecretKeySpec(ByteUtils.mergeByteData(bArr, bArr2), "AES");
```

Figure 2: Decompiled Java method generating an AES session key in version 6.5.0.2170.

https://arxiv.org/pdf/1802.03367.pdf

How do we measure “information”?

- Entropy
  - Don’t be confused if you’ve heard this term in a physics class
    - Entropy in physics is the information we don’t have about energy, which leads to wasted energy
    - Entropy in information theory is a measure of how surprised we’ll be when we learn information, which leads to useful information

Requirements (Shannon, 1948)

1) $I(p) \geq 0$ (information is non-negative, $p \geq 1$)
2) $I(1) = 0$ (events that always occur carry no information)
3) $I(p_1 p_2) = I(p_1) + I(p_2)$ (information due to independent events is additive)

Also, continuity, symmetry, and maximum when all possible events are equiprobable.

$$I(p) = \log(1/p)$$

1) $I(1/2) = 0.30102999566...$
2) $I(1) = 0$
3) $I(1/2) + I(1/3) = \log(2) + \log(3) = 0.77815125038...$

Joint probability: $I(1/6) = 0.77815125038...$

Continuity: $I(1/2.01) = 0.30319605742...$

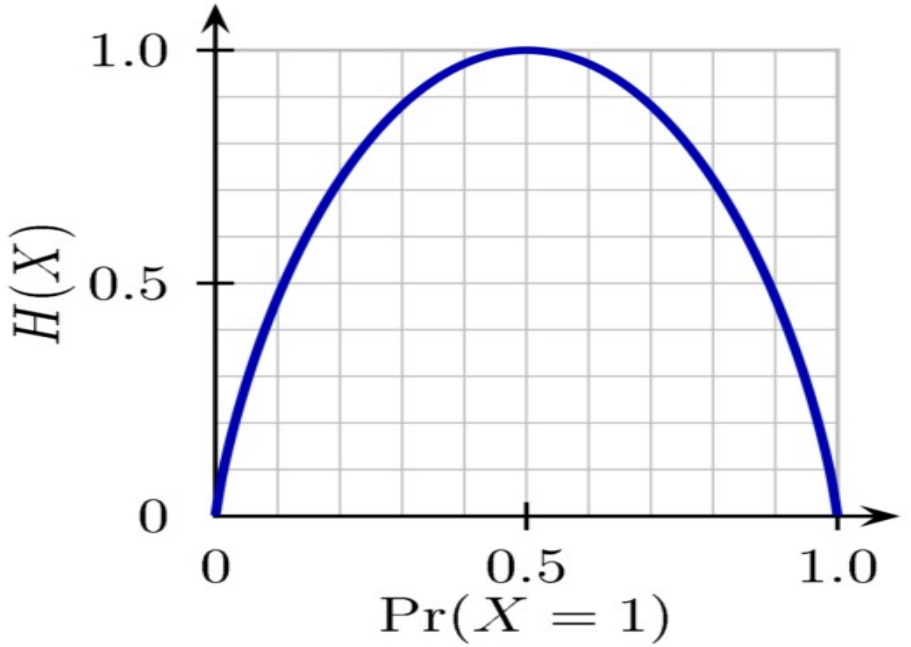
Symmetry: $\log(3) + \log(2) = 0.77815125038...$

Maximum: $\log 3 + \log 3 + \log 3 = 1.43136376416...$

$\log 2 + \log 4 + \log 4 = 1.20411998266...$

Information = Entropy = Surprise

$$H[p] = -\sum_{i=1}^{k} p_i \log p_i$$



Side note – Differential Entropy

https://en.wikipedia.org/wiki/Differential_entropy

$$h(X) = \mathbf{E}[-\log(f(X))] = -\int_{\mathcal{X}} f(x)\log f(x)\,dx$$

f(x) is a probability density function for the signal, the more the signal “jumps around” the higher the entropy, therefore modulating higher frequencies
means more entropy and therefore more bandwidth.

Not a pop quiz #1

- When a 3yo walks by with a stepstool…
  - 4 times out of 10 it’s to get something they’re not supposed to have
  - 2 times out of 10 it’s to climb up to somewhere they’re not supposed to be
  - 1 time out of 10 it’s to wash their hands
  - 1 time out of 10 it’s to get something they’re allowed to have
  - 1 time out of 10 it’s to use as a dollhouse
  - 1 time out of 10 it’s to turn over and use as a storage bin
- What is the entropy of each instance of 3yo stepstool habits?

Answer

Input:

$$-0.4 \log_2(0.4) - 0.2 \log_2(0.2) - 4\,(0.1 \log_2(0.1))$$

Result:

$$2.32193\ldots$$

Not a pop quiz #2

- There are three possible states the Tempe weather could be in during any given hour on a summer day (very hot and bright out, very hot and it’s nighttime, monsoonal rains). What probability distribution over these events would give the maximum entropy in terms of what you might observe in a randomly chosen hour from the summer?

Answer

Input:

$$-\frac{1}{3}\log_2\left(\frac{1}{3}\right)-\frac{1}{3}\log_2\left(\frac{1}{3}\right)-\frac{1}{3}\log_2\left(\frac{1}{3}\right)$$

Exact result:

$$\frac{\log(3)}{\log(2)}$$

$\log(x)$ is the natural logarithm

Decimal approximation:

More digits

1.584962500721156181453738943947816508759814407692481060455...

Not a pop quiz #3

You have 12 coins, one is counterfeit. The counterfeit is either slightly heavier or slightly lighter, otherwise it’s impossible to tell. You have a balance. Using the balance the fewest number of times, find the counterfeit coin.

Not a pop quiz #4

- A measure of the information of a **random process**
- Pop quiz #3: Based on the above definition, order the following binary sequences from most entropy to least entropy?:

  A) 11111111000000000000000011111111

  B) 10111001110011001100010000110100

  C) 00000000000000000000000000000000

  D) 00010010000000100000000100000001

I pity the fool who uses the word “entropy” to describe a bit sequence or string without realizing that they are implicitly talking about algorithmic entropy (*a.k.a.*, Kolmogorov complexity) rather than the standard definition of entropy that Claude Shannon used to describe random processes!

So, where do random numbers for crypto come from in practice on everyday user devices?

Linux kernel's entropy pool

Secure hash function (e.g., SHA1 or ChaCha20-based)

`/dev/urandom`
`/dev/random`