Quantifying Randomness
Part 2: Understanding Entropy

INFO-1301, Quantitative Reasoning 1

University of Colorado Boulder

October 5, 2016

Prof. Michael Paul

Prof. William Aspray

How uncertain is a distribution?

One extreme: everything is equally likely

$P(X=1) = 0.2$
$P(X=2) = 0.2$
$P(X=3) = 0.2$
$P(X=4) = 0.2$
$P(X=5) = 0.2$

With this distribution, you are completely uncertain about what the outcome will be

How uncertain is a distribution?

Another extreme: only one outcome is likely

$P(X=1) = 0.0$
$P(X=2) = 0.0$
$P(X=3) = 1.0$
$P(X=4) = 0.0$
$P(X=5) = 0.0$

With this distribution, you are completely certain about what the outcome will be

Information Entropy

Entropy is a measurement of how evenly distributed a probability distribution is

Lower entropy means it is less even, more certain
Higher entropy means it is more even, less certain

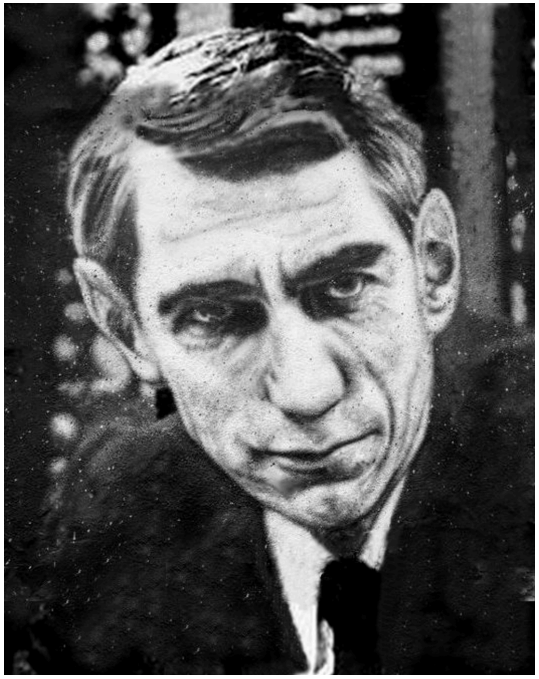
Where did entropy come from?

Entropy is a fundamental part of a discipline of study called **information theory**

Information theory originated in research in telecommunications

- How is information stored?
- How is information transmitted?

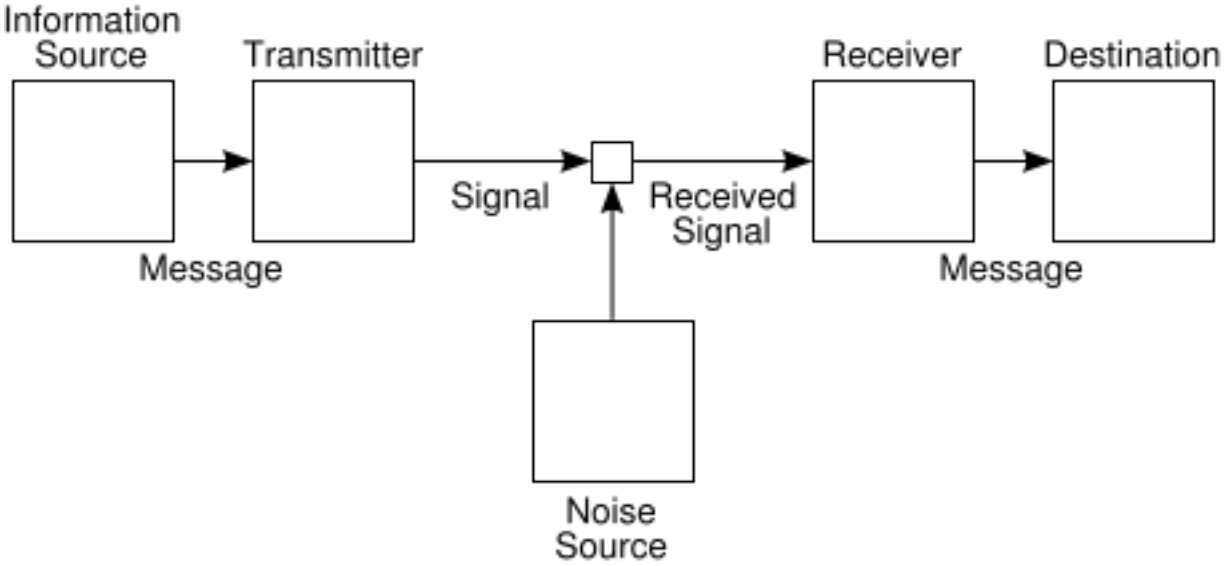
Relatively new insight:
How can we *quantify* information?



Claude Shannon, 1916-2001

Where did entropy come from?

A Mathematical Theory of Communication
by Claude Shannon, 1948



Interpreting Entropy

Which of these variables has more **information?**

P(*X*=1) = 0.0
P(*X*=2) = 0.0
P(*X*=3) = 1.0
P(*X*=4) = 0.0
P(*X*=5) = 0.0

P(*X*=1) = 0.2
P(*X*=2) = 0.2
P(*X*=3) = 0.2
P(*X*=4) = 0.2
P(*X*=5) = 0.2

If I tell you that *X*=3, I didn't tell you anything you didn't already know

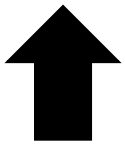
- No new information

Interpreting Entropy

Which of these variables has more **information?**

P(*X*=1) = 0.0
P(*X*=2) = 0.0
P(*X*=3) = 1.0
P(*X*=4) = 0.0
P(*X*=5) = 0.0

P(*X*=1) = 0.2
P(*X*=2) = 0.2
P(*X*=3) = 0.2
P(*X*=4) = 0.2
P(*X*=5) = 0.2



You don't know anything about what the value of *X* might be

- Telling you *X* gives new information

Interpreting Entropy

Entropy is the average number of times you'll be wrong if you guess the answer based on probability

P(*X*=1) = 0.0
P(*X*=2) = 0.0
P(*X*=3) = 1.0 ⬅
P(*X*=4) = 0.0
P(*X*=5) = 0.0

Always guess *X*=3.
Never wrong! So entropy is 0.

Interpreting Entropy

Entropy is the average number of times you'll be wrong if you guess the answer based on probability

$P(X=1) = 0.2$
$P(X=2) = 0.2$
$P(X=3) = 0.2$
$P(X=4) = 0.2$
$P(X=5) = 0.2$

Not clear what to guess first.

Interpreting Entropy

Entropy is the average number of times you'll be wrong if you guess the answer based on probability

$P(X=1) = 0.2$
$P(X=2) = 0.2$
$P(X=3) = 0.2$
$P(X=4) = 0.2$
$P(X=5) = 0.2$

Start with $X=1$.
Wrong 80% of the time.

Interpreting Entropy

Entropy is the average number of times you'll be wrong if you guess the answer based on probability

$P(X=1) = 0.2$
$P(X=2) = 0.2$ ⟵
$P(X=3) = 0.2$
$P(X=4) = 0.2$
$P(X=5) = 0.2$

Move on to $X=2$.
Wrong 75% of the time.

Interpreting Entropy

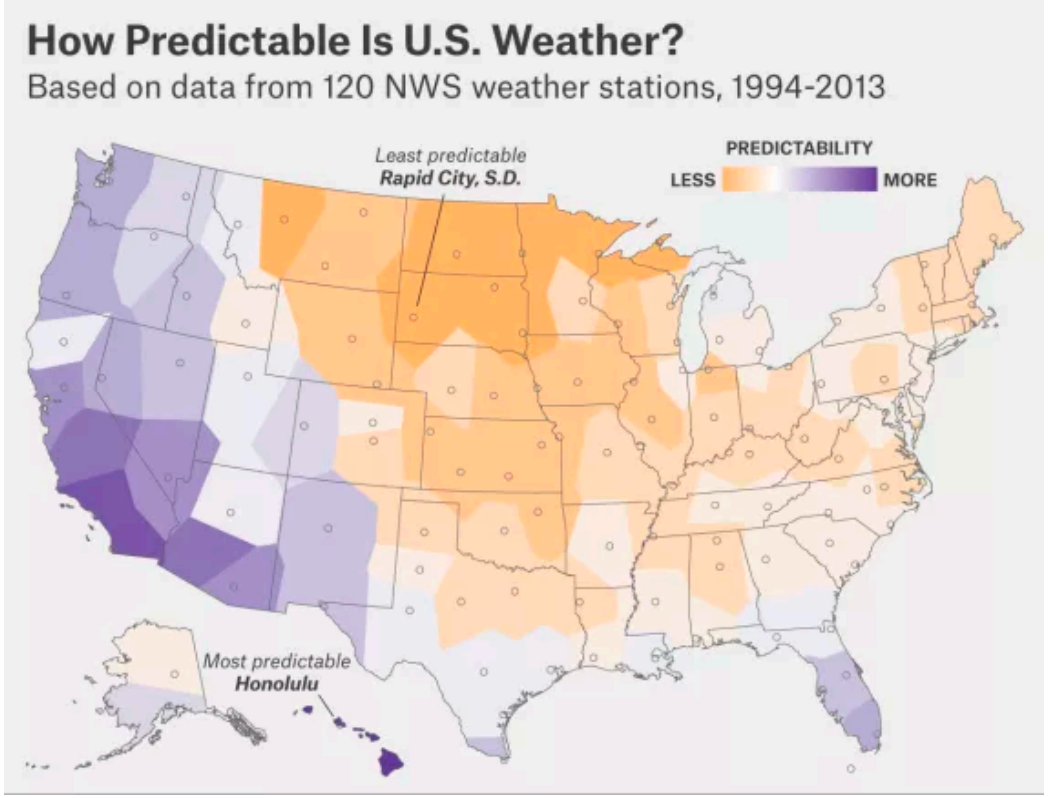
Entropy is the average number of times you'll be wrong if you guess the answer based on probability

$P(X=1) = 0.2$
$P(X=2) = 0.2$
$P(X=3) = 0.2$ ⟵
$P(X=4) = 0.2$
$P(X=5) = 0.2$

Keep repeating until you get the right answer.
On average, you'll have to guess $H(X) = 2.3$ times

Using Entropy

Entropy measures **predictability**



Source: http://fivethirtyeight.com/features/which-city-has-the-most-unpredictable-weather/

Using Entropy

Entropy measures **predictability**

Entropy can be used as a measurement of *risk*, e.g., selecting a stock portfolio

Using Entropy

Entropy measures **equality**

Entropy can measure income equality

- You saw this in your homework

Entropy can measure diversity in a population

- You'll see this today in MiniTab

The more *equal* or *even* a distribution is,
the harder it is to *predict* the outcome