Discrete Maths

25

Information

P.S. Dice || Global Terminology.

{1,4} are termed red outcome.

{2,3,5,6} are termed black outcome.

Rolling a dice

Event A

"it's 4"

$P(A) = 1/6 = 0.1666$

Event B

"it's red"

$P(B) = 2/6 = 1/3 = 0.3333$

$P(A) < P(B)$

$1/6 \,||\, 0.1666$ $\qquad$ $\frac{2}{6}$ $\quad$ 0.3333

Less probability
More Information

more probability
Less Information

Probability in inverse of ~~'Lexical'~~ information.

i.e.

~~Regularity~~ there is ~~no sudden~~ break
at XYZ college,
that is lectures/labs are conducted
100% other than public holidays.

And suddenly somebody told,
"There are no lectures today
(assume it's ~~not public~~ holiday)"

For the real world,
probability of having NO no lectures on regular
day is a little
the information ~~is high~~.

Fraud Detection in Data Analytics

Routinely you are using your credit card in certain season X, but suddenly someone has used the same CC in region Y

there is very less probability/chance but banks consider this as "high value information" to act against fraud.

This discussion shows any unerring ~~message~~ contains information and this can be measured with maths upto a level.

How about formula

$$\text{information} = -\log_2 P(\text{event})$$

it's 4

$P(A) \quad 1/6$

$$-\log_2 1/6$$

$$-(\log_2 1 - \log_2 6)$$

$$= -(-\log_2 6) = 2.585$$

it's red

$2/6$

$$-\log_2 1/3$$

$$-(\log_2 1 - \log_2 3)$$

$$-(-\log_2 3)$$

$$= 1.585$$

Note that information can ~~take greater~~ than 1 value too.

Event A | Event B

it's 6 | it's ~~sed~~ { ~~Can information take -ve value?~~ }

2.585 > 1.585

Event A has given more information compare to event B and math shows that.

Here, when we used log base 2, the measurement unit of information has a global name "bit" binary digit.

Note that $0 \leq p$ is always ≤ 1, ~~$\log_2 p$ is always~~ nonpositive. $-\log_2 p$ is always non-negative.

Moreover,

the smaller the p, larger the quantity $-\log_2 p$ (information).

This is exactly ~~what we want~~.

Other ways are possible

$1/p$, $1/p^2$, $1-p$, $\log_e p$, etc.

Entropy (randomness, chaos, disorder, uncertainty, unpredictability, etc)

"The world and especially life has the highest entropy of them all"

~~entropy of~~ a random variable is a function which attempts to characterize the unpredictability/chaos/disorder/uncertainty of a random variable.

Event A

fair 6 sided dice

possible outcomes are 6

Event B

roulette wheel at Casino

36+0 || 36+0+00

|| Possible outcomes are 36 || ignore, 0/00

which event has more uncertainty?

event → B has more uncertainty and hence more entropy.

"less predictable"

Moreover, entropy is not just about possible outcome but also about their frequency

Event M

Fair die

Event N

Weighted die

i.e.

90% of time it's 2

10% 1, 3, 4, 5, 6

Less uncertainty

less entropy

"more predictable"

Entropy

$$H(X) = -\sum_{x \in X} P(x) \cdot \log P(x)$$

compute entropy of a fair coin.

$$P(X = \text{heads}) = 1/2 \qquad P(X = \text{tails}) = 1/2$$

$$H(P) = -\sum_{x \in (\text{heads}, \text{tails})} P(x) \log P(x)$$

$$= -\left[\tfrac{1}{2} \log \tfrac{1}{2} + \tfrac{1}{2} \log \left(\tfrac{1}{2}\right) \right]$$

$$= -\left[-\tfrac{1}{2} + \left(-\tfrac{1}{2}\right) \right] = -[-1] = 1$$

Notes by Erik G., UMass, Amherst.

Joint entropy

$$H(E, C) = - \sum_{e \in E} \sum_{c \in C} P(e, c) \log(P(e, c))$$

Example

X whether it's sunny or rainy

Y Above ~~70° or below~~ 70°

$P(\text{sunny}, \text{hot}) = 1/2$

$P(\text{sunny}, \text{cool}) = 1/4$

$P(\text{rainy}, \text{hot}) = 1/4$
(Humid)

$P(\text{rainy}, \text{cool}) = 0$, i.e. it's cool when there is thunder or snow.

$$H(x,y) =$$

$$-\left[\frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{4}\log\frac{1}{4} + 0\log 0\right]$$

$$= -\left[-1/2 + -1/2 + -1/2 + 0\right]$$

$$= 3/2$$

~~How about~~ mutual information

"A quantity that measures a relationship between two random variables that are sampled simultaneously"

On average, how much information is communicated in one random variable about another.

We will ~~prove that~~
mutual information both ways is exactly same. $I(A,B) \,||\, I(B,A)$

$$I(X,Y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) \cdot \log\left(\frac{P(x,y)}{P(x)\cdot P(y)}\right)$$

~~Here,~~ $P(X)$ ~~and~~ $P(Y)$

are marginal distributions of X and Y obtained through the marginalization process in probability

"Summing out".

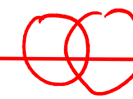
1.5. Independent events

$$P(A \cap B) = P(A) \times P(B)$$

M. Exclusive

$$P(A \cup B) = P(A) + P(B)$$

Not M.E



$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Conditional Prob (Dependent events)

$$P(A \cap B) = P(B) \cdot P(A/B)$$

D.M. Ong C.L. Liu

Mutual Information

$I(A, B)$

Amount of information concerning the occurrence of A that is contained in statement asserting the occurrence of B.

Surprisingly, this is exactly same as

$I(B, A)$

Amount of information concerning the occurrence of B that is contained in statement asserting the occurrence of A.

$$I(A,B) = -\log_2 P(A) - \left(-\log_2 P(A/B)\right)$$

$$= -\log_2 P(A) + \log_2 P(A/B)$$

$$= -\log_2 P(A) + \log_2 \frac{P(A\cap B)}{P(B)}$$

$$= -\log_2 P(A) + \log_2 P(B\cap A) - \log_2 P(B)$$

$$= -\log_2 P(B) + \log_2 P(B\cap A) - \log_2 P(A)$$

$$= -\log_2 P(B) + \log_2 \frac{P(B\cap A)}{P(A)}$$

$$= -\log_2 P(B) + \log_2 P(B/A)$$

$$= -\log_2 P(B) - \left(-\log_2 P(B/A)\right)$$

$$= I(B,A)$$

Note that mutual information can have ~~-ve value~~ to. $x-y$, what if y is greater.

Example

A it's 4

$P(A) = 1/6$

B, it's red $\{1, 4\}$

$P(B) = \frac{2}{6} = 1/3$

C, even number $\{2, 4, 6\}$

$P(C) = \frac{3}{6} = 1/2$

$P(A \cap B) = 1/6 = P(B \cap A)$

$P(A \cap C) = 1/6$

$P(A/B) = \frac{1/6}{1/3} = \frac{1}{\cancel{6}} \times \frac{\cancel{3}}{1} = \frac{1}{2}$

$P(A/C) = \frac{1/6}{1/2} = \frac{1}{\cancel{6}} \times \frac{\cancel{2}}{1} = \frac{1}{3}$

$P(B/A) = \frac{1/6}{1/6} = 1$

$P(C/A) = \frac{1/6}{1/6} = 1$

$$I(A,B) = -\lg P(A) + \lg P(A/B)$$

$$= -\lg \tfrac{1}{6} + \lg \tfrac{1}{2}$$

$$= 2.585 - 1$$

$$= 1.585 \text{ bits}$$

$$I(B,A) = -\lg P(B) + \lg P(B/A)$$

$$= -\lg \left(\tfrac{1}{3}\right) + \lg 1$$

$$= 1.585 + 0$$

$$= 1.585$$

$$I(A, C) = -\log P(A) + \log P(A/C)$$

$$= -\log 1/6 + \log 1/3$$

$$= 2.585 - 1.585$$

$$= 1 \text{ bit}$$

$$I(C, A) = -\log P(C) + \log P(C/A)$$

$$= -\log 1/2 + \log 1$$

$$= -(-1) + 0$$

$$= 1 \text{ bit}$$

Note that M.I. is symmetric measure.

Example

Consider the problem of estimating the likelyhood that there will be 1-hour examination when the professor is scheduled to go out of town.

[observe probability and information, mutual info values carefully]

x_1 : Professor out of town and exam going conducted

x_2 : Prof out of town and exam not conducted

x_3 : Prof in town and exam conducted ???

x_4 : Prof in town and exam not conducted

Let $S = \{x_1, x_2, x_3, x_4\}$

Let A denote the event that an exam is conducted.

B Professor is out of town

$$P(A) = \frac{1}{2} + \frac{3}{16}$$

$$P(A/B) = \frac{1/2}{1/2 + 1/16} = \frac{8}{9}$$

	Prof Out of Town	Exam cond
$P(x_1) = 1/2$	Yes	Yes
$P(x_2) = 1/16$	Yes	No
$P(x_3) = 3/16$	No	Yes
$P(x_4) = 1/4$	No	No

Given

Let C denote the event that ~~the professor~~ is in town

$$P(A/C) = \frac{3/16}{3/16 + 1/4} = \frac{3}{7}$$

Information the exam will be given / already given

$$-\lg P(A) = -\lg \frac{11}{16} = -(\lg 11 - \lg 16) = 3.46 + 4$$
$$= 0.54 \text{ bits}$$

mutual Information about the fact "Professor is out of town" on the fact that exam will be given

$$I(A,B) = -\lg \frac{11}{16} + \lg \frac{8}{9}$$
$$= -3.46 + 4 + 3 - 3.17$$
$$= 0.37 \text{ bits}$$

$$I(A,C) = -\lg \frac{11}{16} + \lg 3/7$$
$$= -3.46 + 4 + 1.58 - 2.81 = -0.69 \text{ bits}$$

Logically, if the prof in town he shall conduct lecture but exam was conducted is surprising shown by $-ve$ MI.

Applications

Communication theory

Transmitter ———— receiver

? PDF ? - Probability density function
(continuous)

~~CDF~~ Cumulative density function

$CDF(x) = Pro(X \le x)$

PMF Probability mass function
(discrete)

Probability distribution function ?

- funfair rings and products play
- shoot balloons
- ~~what shall~~ be MRP of a product ?
- why casino / lottery company will always have profit ?
- MCQ
- Bernoulli trials
- Movies

  Focus 2015

  Jerry and Marge go large

  Teen Patti

  21 tablet

  Travelling Sale Prob. (TSP)