INFORMATION THEORY

and

CODING

Husam Abduldaem Mohammed

Lecture two

probability

Suppose that one of the possible outcomes of an experiment is called A and that when the experiment is repeated N times the outcome A occurs N_A times.

The relative frequency of the occurrence of A is N_A / N.

the limiting value of the relative frequency of occurrence is called the probability of outcomes A, written p(A), so that

$$p(A) = \lim_{N \longrightarrow \infty} N_A / N$$

An alternative definition of probability of occurrence of an event A is

$$p(A) = \frac{number\ \ of\ \ possible\ \ favorable\ \ outcomes}{total\ \ number\ \ of\ \ possible\ \ equally\ \ likely\ \ outcomes}$$

it is apparent from either definition that the probability of occurrence of an event p is a positive number and that

$$0 \le P \le 1$$

If an event is not possible then **p = 0**, while if an event is certain then **p = 1.**

Mutually exclusive events

Two possible outcomes of an experiment are defined as being mutualy exclusive if the occurrence of one outcome precludes the occurrence of the other.

For suppose that in a very large number N of repetitions of the experiment, outcome $\mathbf{A_1}$ had occurred $\mathbf{N_1}$ times and outcome $\mathbf{A_2}$ had occurred $\mathbf{N_2}$ times. Then, $\mathbf{A_1}$ or $\mathbf{A_2}$ will have occurred $\mathbf{N_1 + N_2}$ times and

$$p(A_1 \text{ or } A_2) = \frac{N_1 + N_2}{N} = \frac{N_1}{N} + \frac{N_2}{N}$$

$$p(A_1 \text{ or } A_2) = p(A_1) + p(A_2) \qquad [1]$$

Equation (1) may be extended to more than two mutually exclusive outcomes thus:

$$p(A_1 \text{ or } A_2 \text{ or } \ldots\ldots\ldots \text{ or } A_L) = \sum_{j=1}^{L} p(A_j)$$

Example 1.

Suppose two coins are tossed together there are four mutually exclusive and equally likely outcomes HH, HT, TH,and TT. Since the number of equally likely outcomes is M = 4. The probability of any one is 0.25.

p(HH) = p(HT) = p(TH) = p(TT) = 0.25.

P(not HH) = 1- 0.25 = 0.75.

P(match) = p(HH) + p(TT) = 0.25 + 0.25 = 0.5.

Example 2.

Suppose that we have 10 balls in a box. The balls are identical except that 8 are white and 2 are black. What is the probability that in a single draw we shall select the ball.

solution

$$p(B) = \frac{2}{10} = \frac{1}{5}$$

$$p(W) = \frac{8}{10} = \frac{4}{5}$$

$$p(W) = p(W \ or \ B) - p(B) = 1 - \frac{1}{5} = \frac{4}{5}$$

Joint and conditional probabilities

consider the events **A** and **B** which may or may not occur together and let the joint probability be **p (AB).** We repeat the experiment **N** times where **N** is very large and record $\mathbf{N_{AB}}$ the number of times A and B occur together. The joint probability is

$$\text{p (AB)} = \frac{\text{N}_{\text{AB}}}{\text{N}}$$

In this process **A** has occurred $\mathbf{N_A}$ times with or without **B**, so $\mathbf{N_A}$ includes $\mathbf{N_{AB}}$ and

$$N_A \geq N_{AB}$$

It is quit possible that the occurrence of **B** depends in some way on the occurrence of **A**. conditional probability **p (B/A)** is the probability of **B** given **A** has occurred.

$$p\left(B/A\right) = \frac{N_{AB}}{N_A} = \frac{\frac{N_{AB}}{N}}{\frac{N_A}{N}} = \frac{p(AB)}{p(A)}$$

$$\text{so} \qquad p(AB) = p\left(B/A\right)p(A)$$

- From the Product Rule and the symmetry that p (AB) = p(BA), it is clear that

p(A/B)p(B) = p(B/A)p(A). **Bayes' Theorem** then follows:

Bayes' Rule:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}$$

- The importance of Bayes Rule is that it allows us to reverse the conditionalizing of events, and to compute p(B/A) from knowledge of p(A/B); p(A); and p(B): Often these are expressed as prior and posterior probabilities, or as the conditionalizing of hypotheses upon data.

Statistical independent events

Suppose the event **B** is independent of **A** so the occurrence of **A** does not influence the occurrence of **B**.

$$p(B/A) = p(B)$$

Hence for statistically independent event

$$p(AB) = p(A)p(B)$$

Example1:

Two boxes contain white and black balls. Box A contains 2 black balls and 1 white ball. Box B contains 3 black balls and 2 white balls. One of the boxes is selected at random and one of the balls is chosen. What is the probability of drawing a white ball.

Solution:

$p(W) = p(c_1W) \; + p(c_2W)$

$p(c_1W) = p(c_1)\, p(W/c_1) = 1/2\ (1/3) =1/6$

$p(c_2W) = p(c_2)\, p(W/c_2)\ 1/2\ (2/5) =1/5$

$p(W) =1/6\ (1/5) =11/30$

Example 2:

a box containing two white balls and three black balls. Two balls are drawn in succession. If the first one not been relaced, what is the chance of drawing two white balls in succession.

Solution:

Letting event A represents a white ball in draw 1.

Event B represents a white ball on draw 2.

P(A) = 2/5

p(B/A) = 1/4

so

p(AB) = p(A) p(B/A) = (2/5) (1/4) = 2/20 =1/10

if the first ball is replaced then

p(AB) = (2/5)(2/5) = 4/25

Statistical average

assume that we have a discrete variable x whose average value we would like to calculate.

x can takes over the values x_1, x_2, x_3, ……, x_n. We perform repetition of experiments measuring the number of times N_1 that x_1 appears, the number N_2 that x_2 appears and so on.

The total number of trials N is $N_1 + N_2 + N_3 + \ldots\ldots\ldots + N_n = N$

$$\mathrm{av}(x) = \frac{xN_1 + xN_2 + xN_3 + \ldots\ldots + xN_n}{N}$$

$$= x_1 \frac{N_1}{N} + x_2 \frac{N_2}{N} + x_3 \frac{N_3}{N} + \ldots . + x_n \frac{N_n}{N}$$

$$\mathrm{av}(x) = \sum_{i=1}^{n} x_i p(x_i) = m_1$$

The mean square value can be calculated as in below

$$\mathrm{av}(x^2) = \sum_{i=1}^{n} x_i^2 p(x_i) = m_2$$

and

$$\text{variance} = \sum_{i=1}^{n} (x_i - m_1)^2 p(x_i) = \sigma^2$$

$$\text{standard deviation} = \sqrt{\text{variance}} = \sigma$$

Homework1:

1) suppose that a class of 100 student is given a test and the following scores result

score:	90	85	80	75	70	65	60
no. of students:	1	4	20	50	20	4	1

find the average score.

2) consider the transmission of binary digits through a channel as might occur, for example, in computer system. As is customary, we denote the two possible symbols as 0 and 1. Let the probability of receiving 0, given a 0 was sent, p(0r/0s), and the probability of receiving a 1, given a 1 was send p(1r/1s) be

p(0r/0s) = p(1r/1s) = 0.9, and p(0s)=0.8.

Find the probabilities p(1s/1r), and p(0s/1r), p(0s/or) and p(1s/0r).

3) Jo has a test for a nasty disease. We denote Jo's state of health by the variable a and the test result by b.

a = 1 Jo has the disease

a = 0 Jo does not have the disease

The result of the test is either `positive' (b = 1) or `negative' (b = 0); the test is 95% reliable: in 95% of cases of people who really have the disease, a positive result is returned, and in 95% of cases of people who do not have the disease, a negative result is obtained. The final piece of background information is that 1% of people of Jo's age and background have the disease.

OK { Jo has the test, and the result was positive. What is the probability that Jo has the disease?

Probability density function:-

$$1-\ p(x_1 < x < x_2) = \int_{x_1}^{x_2} f(x)dx$$

$$2-\ f(x) \geq 0\, becouse \quad 0 \leq p \leq 1$$

$$3-\ \int_{-\infty}^{\infty} f(x) = 1$$

If we have continuous variable x then :-

$$av(x) = \int_{-\infty}^{\infty} xf(x)dx = m1$$

$$av(x^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = m2$$

$$\text{var}\, iance = \int_{-\infty}^{\infty} (x - m_1)^2 f(x)dx$$

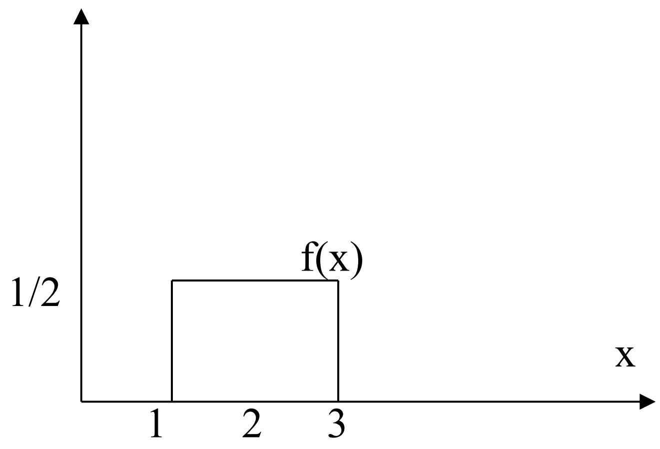
Example1

: For the pdf shown

1.find the average length

2.find the mean square value of x

3.find p(x > 2.5)



Sol:-

$$1.m_1 \int_{-\infty}^{\infty} xf(x)dx$$

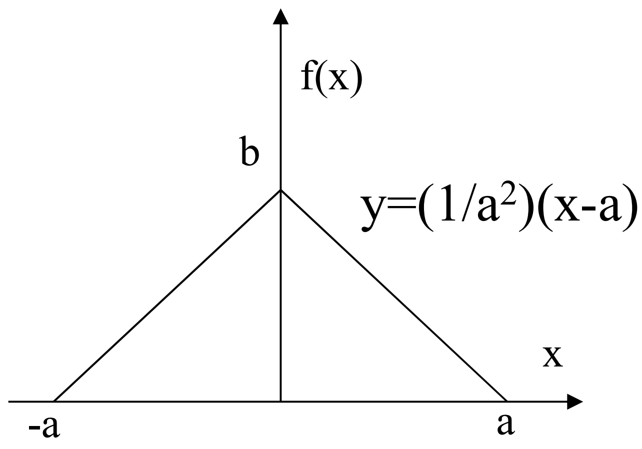
$$= \int_{1}^{3} x(\frac{1}{2})dx = \frac{1}{2}.\frac{x^2}{2} = \frac{1}{2}(\frac{9}{2} - \frac{1}{2}) = 2$$

$$2.m_2 \int_{-\infty}^{\infty} x^2 f(x)dx = \int_{1}^{3} x^2(\frac{1}{2})dx = \frac{1}{2}.\frac{x^3}{3}$$

$$3.p(x > 2.5) = \int_{2.5}^{3} \frac{1}{2}dx = \frac{1}{2}x = \frac{1}{4}$$

Example2:-

1. Find b in terms of a
2. What p(x>9/2)



Sol:-

$$1. Area = 1 \Rightarrow \frac{1}{2} base * high = 1$$

$$\frac{1}{2}(2a)b = 1 \Rightarrow ab = 1 \Rightarrow b = \frac{1}{a}$$

$$2. p(x > \frac{a}{2}) = \int_{a/2}^{a} f(x)dx = \int_{a/2}^{a} -\frac{1}{a^2}(y-a)dx = \int_{a/2}^{a} \frac{-x}{a^2}dx + \int_{a/2}^{a} \frac{1}{a}dx$$

$$= -\frac{1}{a^2}(\frac{a^2}{2} - \frac{a^2}{s}) + \frac{1}{a}(a - \frac{a}{z}) = (\frac{1}{8} - \frac{1}{2}) + (1 - \frac{1}{2}) = -\frac{6}{16} + \frac{1}{2} = \frac{4}{32} = \frac{1}{8}$$

Commulative distribution function (cdf)

The probability distribution function of x is the function

$$F(x) = \int_{-\infty}^{x} f(x)\,dx$$

$$f(x) = \frac{d}{dx} F(x)$$

$$F(x_1) = P(x < x_1) = \int_{-\infty}^{x_1} f(x)\,dx$$

Properties of F(x)

1. $F(-\infty) = 0$
2. $F(\infty) = 1$
3. $0 \le F(x) \le 1$
4. $F(x_1) \le F(x_2)$ if $x_1 < x_2$
5. $P(x_1 < X \le x_2) = F(x_2) - (x_1)$

Example1 :-

1. find p(o < x < ∞)
2. find P(3 > x > 6)

sol :-

$$1.p(o < x < \infty) = \int_{o}^{\infty} f(x)dx$$

$$= \int_{o}^{\infty} \frac{1}{16} dx = 8 * \frac{1}{16} = \frac{1}{2}$$

$$2.p(3 < x < 6) = \int_{3}^{6} f(x)dx = \frac{1}{16}(6-3) = \frac{3}{16}$$

Homework2:-

1-

A. find $f(x)$
B. find $p(x>0.5)$
C. find $p(x>4)$

2-

A. find $E(x) = m1$
B. find $F(x)$
C. find $P(x < 1)$
D. find $P(1 < x < 1.5)$