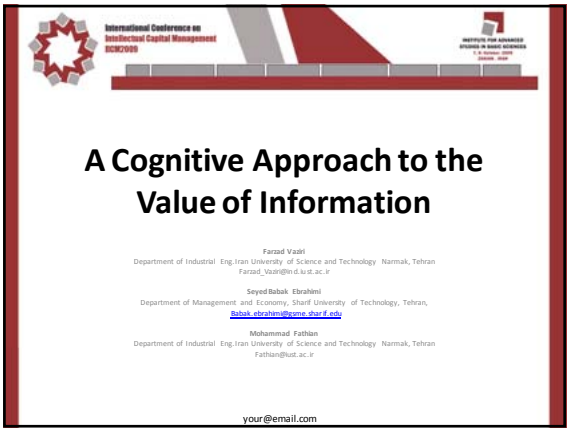


A Cognitive Approach to the Value of Information

Farzad Vaziri
Department of Industrial Eng.Iran University of Science and Technology Narmak, Tehran
Farzad_Vaziri@ind.iust.ac.ir

Seyed Babak Ebrahimi
Department of Management and Economy, Sharif University of Technology, Tehran,
Babak_ebrahimi@gsme.sharif.edu

Mohammad Fathian
Department of Industrial Eng.Iran University of Science and Technology Narmak, Tehran
Fathian@iust.ac.ir

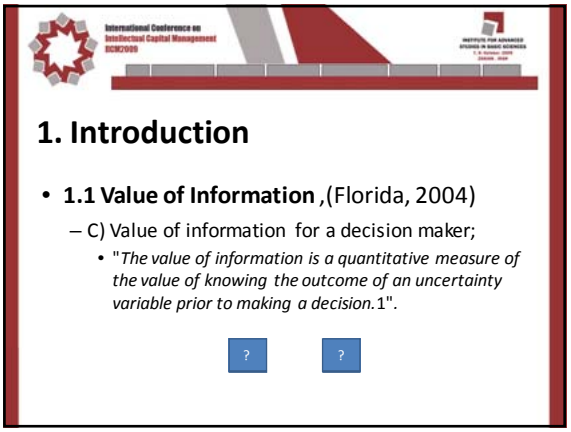
your@email.com

1. Introduction

- **1.1 Value of Information** ,(Florida, 2004)
  - A) Value of information as a product;
    - is what can be pointed in new digital economy and has a very fast increasing usage now a days, information products are what can be supplied into the market as information. These kinds of products are have market value which economical phenomena such as supply and demand have the most effects on them. (Quah, 2003)

1. Introduction

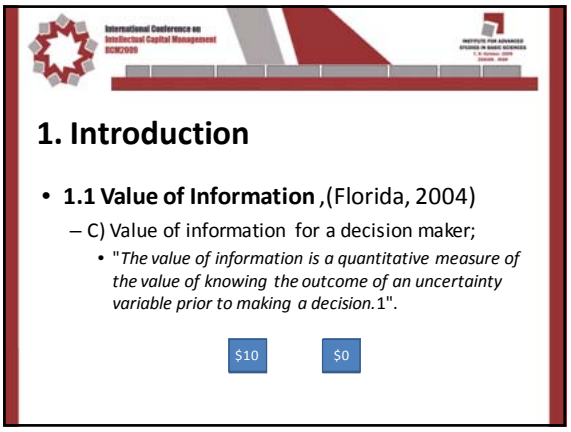
- **1.1 Value of Information** ,(Florida, 2004)
  - B) Operational value of information in information systems;
    - comes from it's usage in information and processing systems. There value of information is equal to its facilitation in processing.(Fenne,2002) and (Powell,1003)
      - i.e. in marketing or in service industry



1. Introduction

- **1.1 Value of Information** ,(Florida, 2004)
  - C) Value of information for a decision maker;
    - "*The value of information is a quantitative measure of the value of knowing the outcome of an uncertainty variable prior to making a decision.*1".

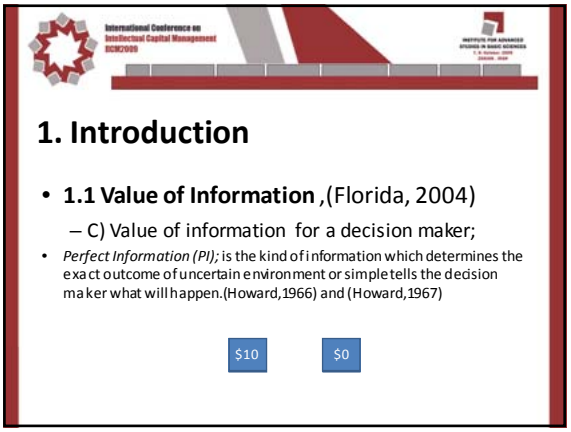
? ?



1. Introduction

- **1.1 Value of Information** ,(Florida, 2004)
  - C) Value of information for a decision maker;
    - "*The value of information is a quantitative measure of the value of knowing the outcome of an uncertainty variable prior to making a decision.*1".

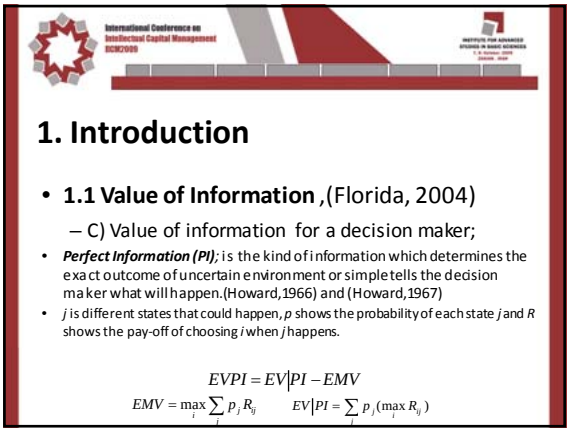
$10 $0



1. Introduction

- **1.1 Value of Information** ,(Florida, 2004)
  - C) Value of information for a decision maker;
- *Perfect Information (PI);* is the kind of information which determines the exact outcome of uncertain environment or simple tells the decision maker what will happen.(Howard,1966) and (Howard,1967)

$10 $0

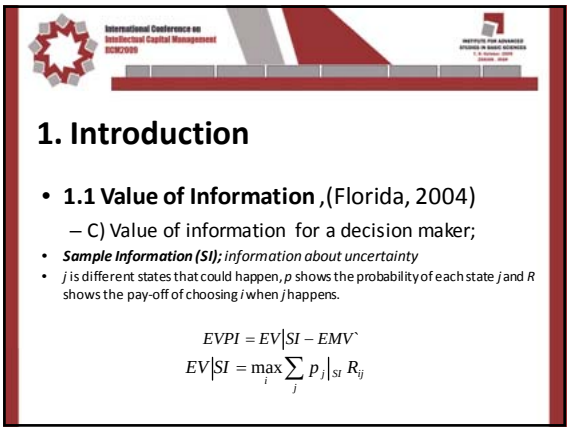


1. Introduction

- **1.1 Value of Information** ,(Florida, 2004)
  - C) Value of information for a decision maker;
- ***Perfect Information (PI);*** is the kind of information which determines the exact outcome of uncertain environment or simple tells the decision maker what will happen.(Howard,1966) and (Howard,1967)
- *j* is different states that could happen, *p* shows the probability of each state *j* and *R* shows the pay-off of choosing *i* when *j* happens.

$$EVPI = EV|PI - EMV$$

$$EMV = \max_i \sum_j p_j R_{ij} \qquad EV|PI = \sum_j p_j (\max_i R_{ij})$$

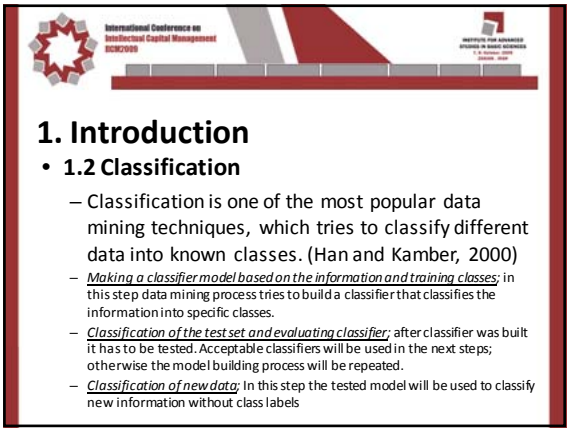


1. Introduction

- **1.1 Value of Information** ,(Florida, 2004)
  - C) Value of information for a decision maker;
- ***Sample Information (SI);*** *information about uncertainty*
- *j* is different states that could happen, *p* shows the probability of each state *j* and *R* shows the pay-off of choosing *i* when *j* happens.

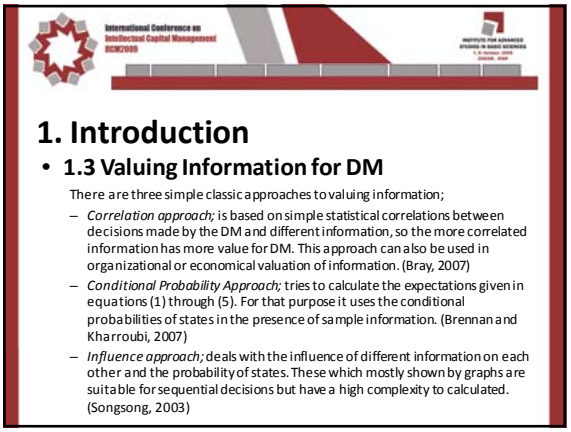
$$EVPI = EV|SI - EMV`$$

$$EV|SI = \max_i \sum_j p_j|_{SI} R_{ij}$$



1. Introduction

- **1.2 Classification**
  - Classification is one of the most popular data mining techniques, which tries to classify different data into known classes. (Han and Kamber, 2000)
  - *Making a classifier model based on the information and training classes;* in this step data mining process tries to build a classifier that classifies the information into specific classes.
  - *Classification of the test set and evaluating classifier;* after classifier was built it has to be tested. Acceptable classifiers will be used in the next steps; otherwise the model building process will be repeated.
  - *Classification of new data;* In this step the tested model will be used to classify new information without class labels

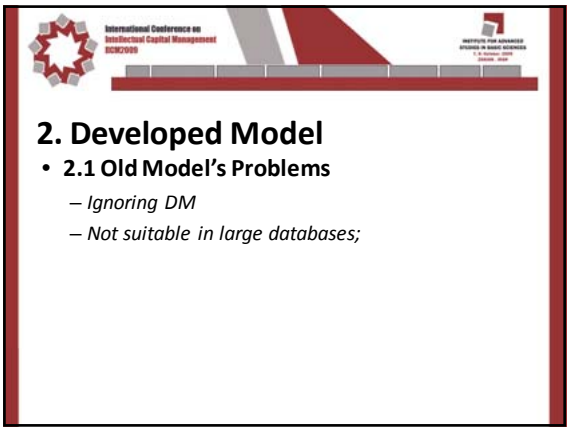


1. Introduction

- **1.3 Valuing Information for DM**

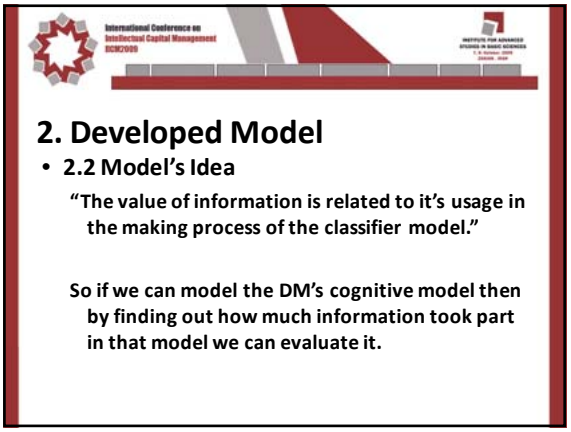
There are three simple classic approaches to valuing information;

- *Correlation approach;* is based on simple statistical correlations between decisions made by the DM and different information, so the more correlated information has more value for DM. This approach can also be used in organizational or economical valuation of information. (Bray, 2007)
- *Conditional Probability Approach;* tries to calculate the expectations given in equations (1) through (5). For that purpose it uses the conditional probabilities of states in the presence of sample information. (Brennan and Kharroubi, 2007)
- *Influence approach;* deals with the influence of different information on each other and the probability of states. These which mostly shown by graphs are suitable for sequential decisions but have a high complexity to calculated. (Songsong, 2003)



2. Developed Model

- **2.1 Old Model's Problems**
  - *Ignoring DM*
  - *Not suitable in large databases;*



2. Developed Model

- **2.2 Model's Idea**

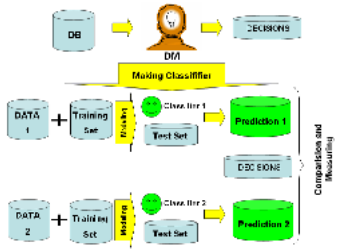
"The value of information is related to it's usage in the making process of the classifier model."

So if we can model the DM's cognitive model then by finding out how much information took part in that model we can evaluate it.

2. Developed Model

- **2.3 Model**
  - M(*t*,*traindb*) is a classifier made by method *t* of classifications based on the train database *traindb*.
  - O(M(*t*,*traindb*),D) which is predictions of classes for each record of the database D.
  - change the O into a single output function if we use each record at a time such as D(*i*), so it would look like, O(M(*t*,*traindb*),D(*i*)),which shows the prediction of class for the *i*th record.
  - Lets abstract the O(M(*t*,*traindb*),D(*i*)) into c(*i*)

2. Developed Model



2. Developed Model

- *Assumption 1.* The classifiers that we use predict only one class, the one with the most probability to occur.
  - function v shows the value gained by the DM in different situations. nc shows the condition when c did not happened and pc shows the prediction of c. in this equation c shows different classes that can be predicted.

$$EV(D(i)) = EV(s(i))$$

$$EV(D(i)) = v(s(i),c(i))\,p(c(i)|pc(i)) + v(s(i),nc(i))\,p(nc(i)|pc(i))$$

$$v(s(i),nc(i))\,p(nc(i)|pc(i)) = \sum_{c \in C, c \neq c(i)} v(s(i),c)\,p(c|pc(i))$$

2. Developed Model

– When we have real databases with known classes we can measure the value of information in the databases. If we show the read class of *i*th record with rc(*i*) then it would be

$$V(d,D)=\sum_{i} v(s(i),rc(i))$$

– Here V shows the value of subset data d in the space of D, for example d can be some columns of the database,table, D.

2. Developed Model

– And the value of information (VOI) would come from the

$$VOI(d,D)=V(d,D)-V(\emptyset,D)$$

$$V(\emptyset,D)=\max_{s}\sum_{i} v(s,rc(i))$$

2. Developed Model

- The utility of DM have to be considered in two places, first in the calculation of v, and then in choosing the optimum strategy s. s in the prediction of pc would come from the

$$s(pc)=s\Big|\max_{s}\sum_{c\in C} v(s,c)\,p(c\,|\,pc)$$

2. Developed Model

- *Assumption 2.*

$$\forall c \in C(if\ (\forall(c' \in C, c' \neq c)p(c) > p(c')) \rightarrow \exists s | v(s,c)p(c) > \sum_{c'} v(s,c')p(c')$$

- With assumption 2 we could eliminate the effect of utilities on the process of choosing strategies. This elimination process means that the strategy that will be chosen in the situation of a pc (prediction of c) is the strategy that maximizes the utility when c really happens.

2. Developed Model

$$VOI(d,D) = n - error(M(t,d),D) - V(0,D)$$