International Business and Economics Research Conference

A Measure of Relevance for Decision Analysis

Roberto Ley-Borrás

Instituto Tecnológico de Orizaba

Mexico

http://decidir.org *Las Vegas, NV, October 7, 2003*

The issues of this talk

- Modeling decisions.
- Representing relevance (probabilistic dependence).
- Closeness and similarity between probability distributions.
- Measuring relevance using closeness.
- Using the measure of relevance.

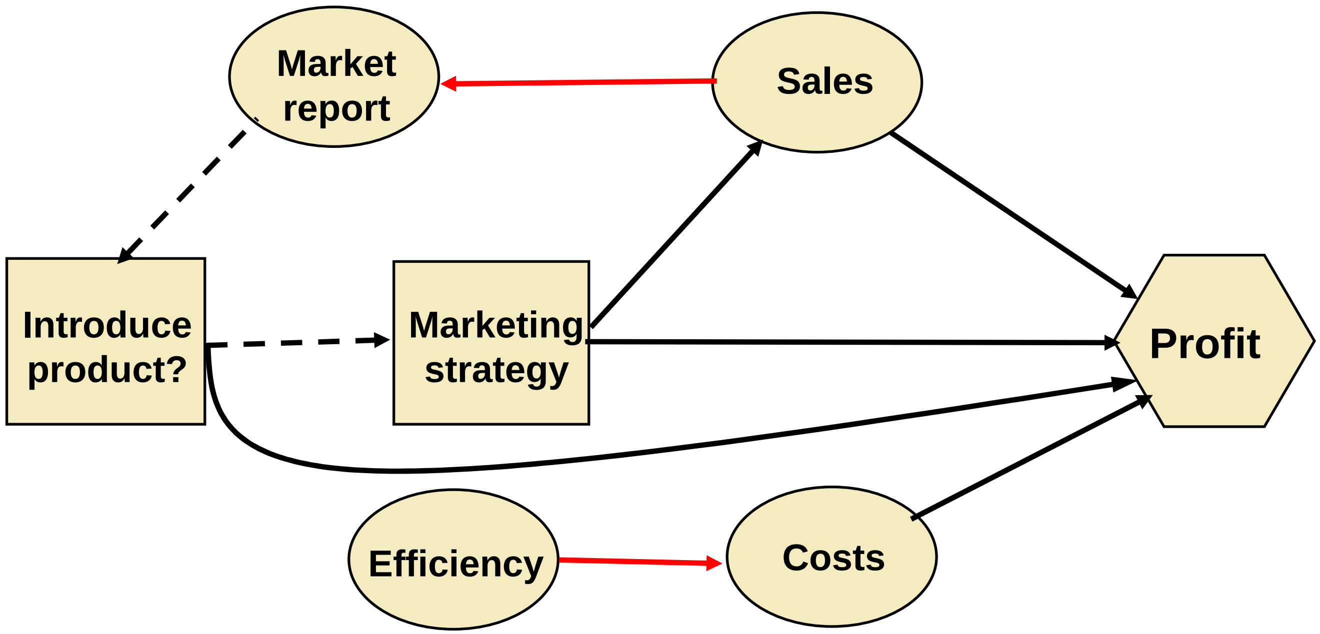
Decision analysis helps tackle difficult problems

- **The objective of decision analysis is to obtain clarity of action.**
- **Decision analysis is specially useful when the situations include:**
  - **complexity,**
  - **high stakes,**
  - **unclear objectives or**
  - **uncertainty.**

Modeling decision helps overcome these difficulties

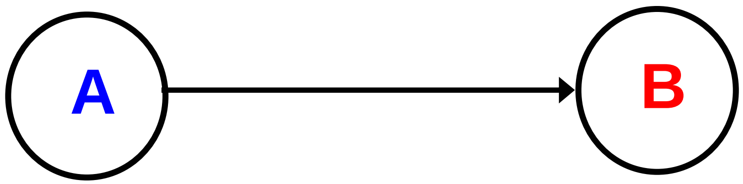
- **A leading modeling tool is the influence diagram.**
- **Influence diagrams are a compact representation of decisions, uncertain events and values, and their relationships.**
- **A most valuable feature of influence diagrams is their ability to model relevance (probabilistic dependence) among uncertain events.**

An Influence Diagram



- **Elements: Decisions, Uncertain events, Values.**
- **Relationships: Information, Relevance, Impact.**

Representing relevance between uncertain events



- **Each node (uncertain event) contains one or more probability distributions of the event.**
- **Relevance is probabilistic dependence.**
- **The arrow between event A and event B means that the probability distribution of B changes for each possible state of A.**
- **The lack of an arrow between two events indicates irrelevance or probabilistic independence.**

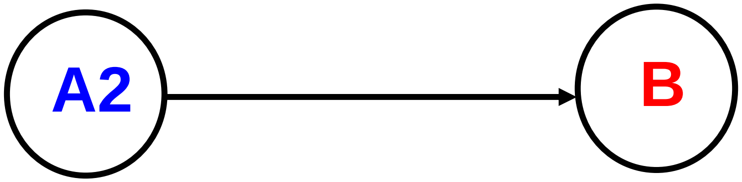
A simple example of irrelevance (probabilistic independence)

- **Event A1: Shift in taste from regular produce to organic growth produce.**
- **Event B: Annual sales of laptops in the domestic market.**

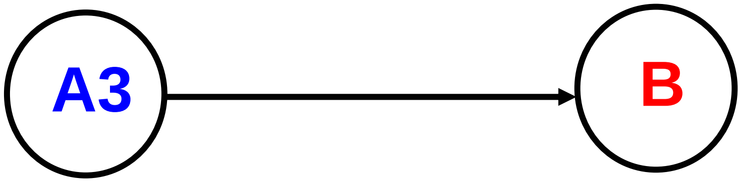
- **The lack of arrow tells us immediately that the probability of laptop sales will be unaffected by whether a shift in taste happens or not.**

Two examples of relevance

- **Event A2: Achieving a 4% increase in Gross National Product.**
- **Event B: Annual sales of laptops in the domestic market.**



- **Event A3: Achieving a technological breakthrough that decreases laptop prices by 40%.**



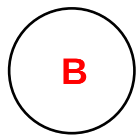
Simply stating relevance does not tell the whole story

- **The current graphical representation of relevance is not very informative because the size of the effect of the relevant variable is not displayed.**
- **A reason for this deficiency is that <u>we do not have an intuitive, easy to compute measure of relevance</u>.**
- **We can improve communication among analysts and clients by measuring relevance.**

Relevance: difference between prior and conditioned probabilities

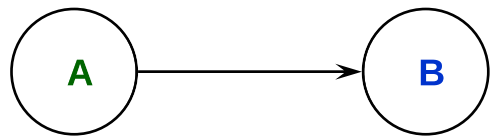
We can recognize relevance in the difference between the prior probability distribution and the conditioned probability distributions (conditioned on each outcome).

Example:



Prior probability

	{B}
B_1	.20
B_2	.50
B_3	.30

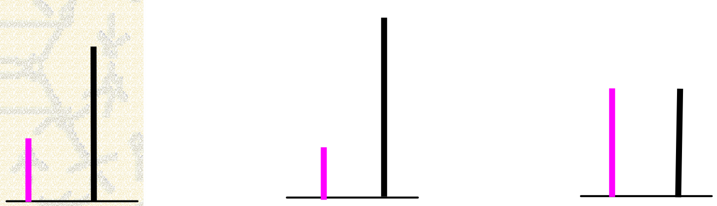


Conditioned probabilities

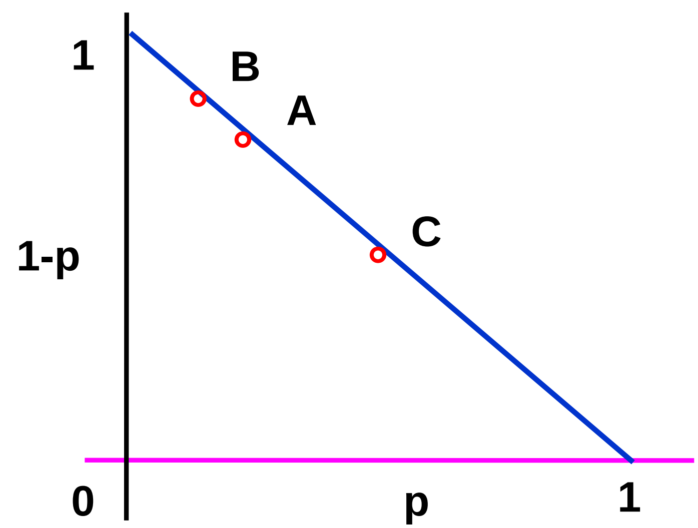
	$\{B\|A_1\}$	$\{B\|A_2\}$	$\{B\|A_3\}$
B_1	.85	.30	.10
B_2	.10	.40	.20
B_3	.05	.30	.70

Similarity (closeness) among probability distributions

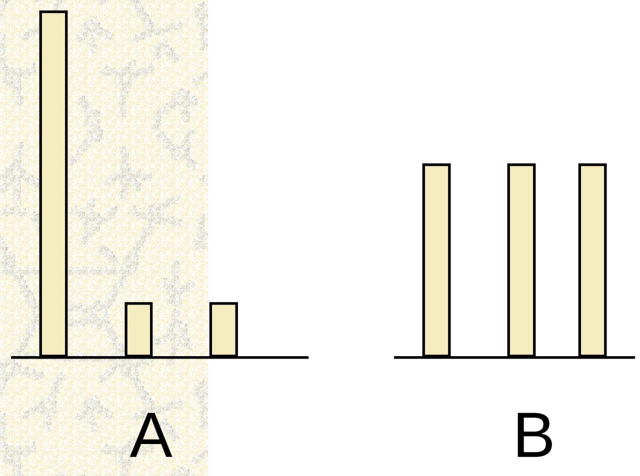
1. **We have an intuitive sense of similarity (closeness) between probability distributions.**
2. **We can tell how close two probability distributions are comparing their corresponding moments.**
3. **We can represent an n-outcome probability distribution as a point in an n-dimensional space.**

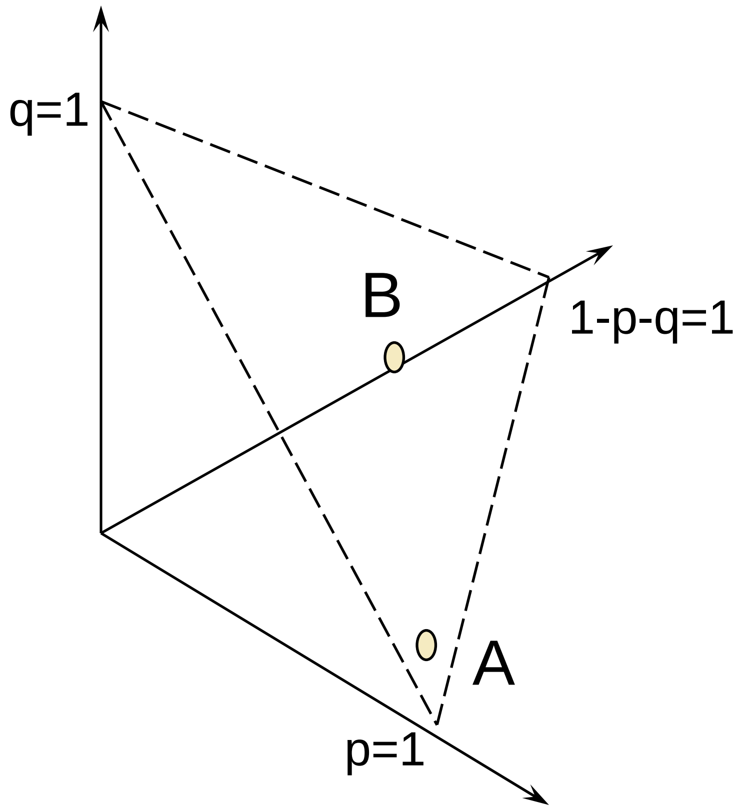


{A} **{B}** **{C}**



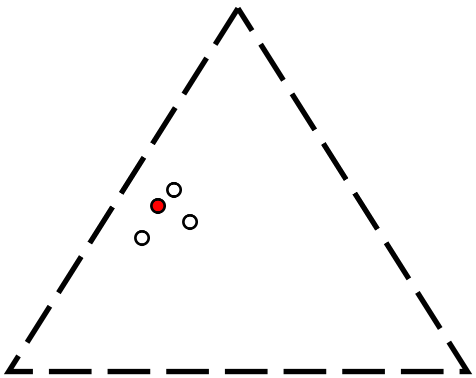
Three-Outcome Probability Distributions

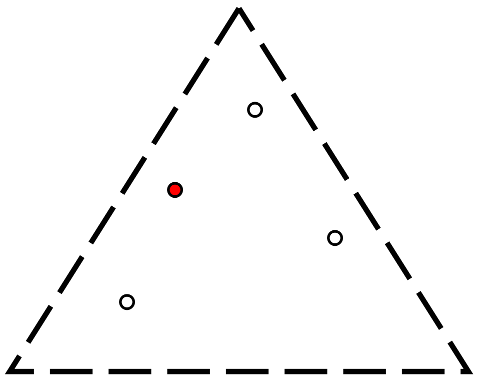




Geometric distance from the prior distribution measures relevance

- The closeness of the points indicates similarity of the corresponding probability distributions.





- The further away a new point is from the point representing a prior probability distribution, the more informative the conditioning outcome is.
- **The larger the distance, the more relevant the outcome**.

Using Geometric Distance to Measure Relevance

***Relevance to event B of state A_i*, R(B, A_i), is defined as the geometric distance between two points, one representing the prior probability distribution of B, {B}, with values p_1, p_2, ..., p_n, and the other representing the posterior probability distribution of B given the state A_i, {B|A_i}, with values q_1, q_2, ..., q_n.**

$$R(B,A_i) = [\ (p_1-q_1)^2 + (p_2-q_2)^2 + \ldots + (p_n-q_n)^2\ ]^{1/2}$$

R(B, A_i) measures the shape difference between the prior probability distribution of B and the posterior probability of B given a particular outcome A_i.

Measuring relevance: An example

	Prior of B {B}		Conditional probability of B given A {B|A₁}	{B|A₂}	{B|A₃}
B_1	.25	B_1	.70	.10	.20
B_2	.40	B_2	.20	.20	.60
B_3	.35	B_3	.10	.70	.20

Relevance of the states of A to event B are:

$R(B,A_1) = 0.552$ $R(B,A_2) = 0.430$ $R(B,A_3)= 0.255$

- **The lowest value of relevance, $R(B,A_3)$, corresponds to the conditioned probability distribution most similar to the prior distribution of A (the third one).**

Weighted Relevance measures the overall effect of the conditioning event

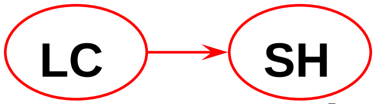
- **Relevance between the prior and each conditioned probability distribution tells us how informative each outcome is.**
- **To measure the overall importance of the probabilistic link (the effect of the conditioning event), we can compute weighted relevance as the expected value of relevance:**

$$\langle R(B,A)\rangle = \sum_{i=1}^{n} \{A_i\}\, R(B,A_i)$$

- **Relevance of each outcome is weighted by the probability of that outcome happening.**

Computing Weighted Relevance: relevance of smoking to lung cancer

Measuring relevance of smoking habit (SH) to lung cancer (LC), and vice versa.



S	.250
S'	.750

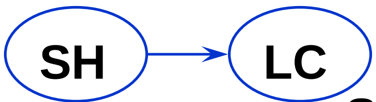
	L	L´
S	.769	.233
S'	.231	.767

Relevance of *Lung cancer* (L) and *No lung cancer* (L'):

R(SH, L) = .7343

R(SH, L') = .0247

Weighted Relevance: <R(SH, LC)> = .0325(.7343) + .9675(.0247) = .0477



L	.0325
L'	.9675

	S	S´
L	.0100	.0100
L'	.9000	.9900

Relevance of *Smoking* (S) and *No smoking* (S'):

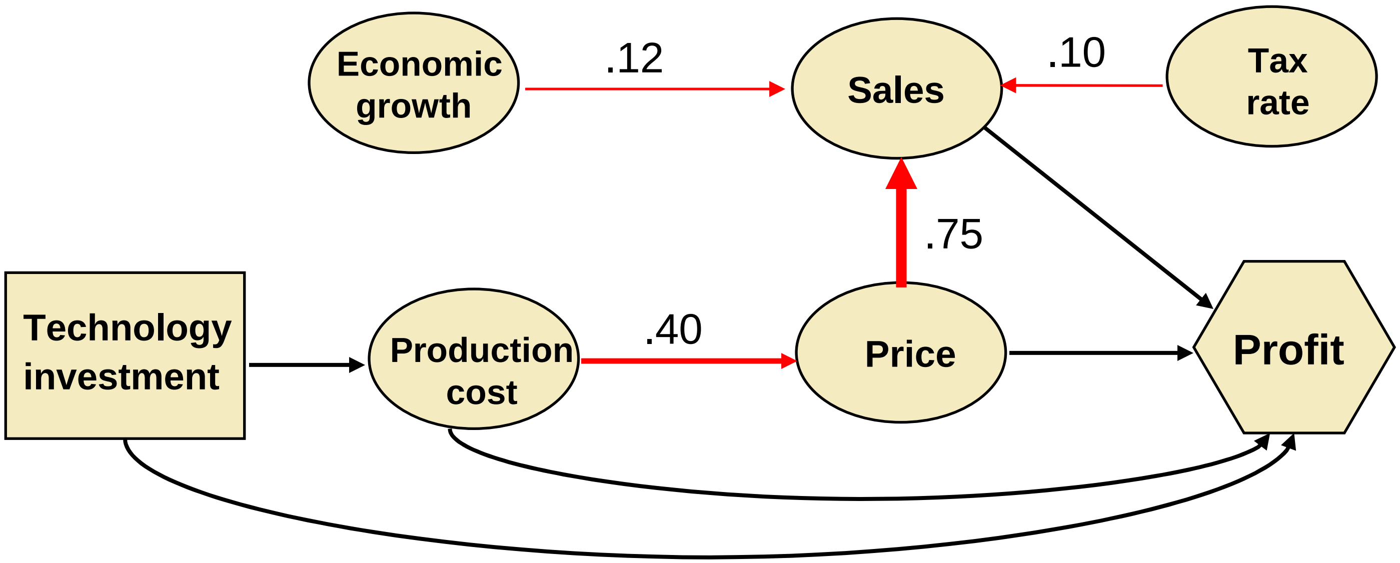
R(LC, S) = .0955

R(LC, S') = .0318

Weighted Relevance: <R(LC, SH)> = .250(.0955) + .750(.0318) = .0477

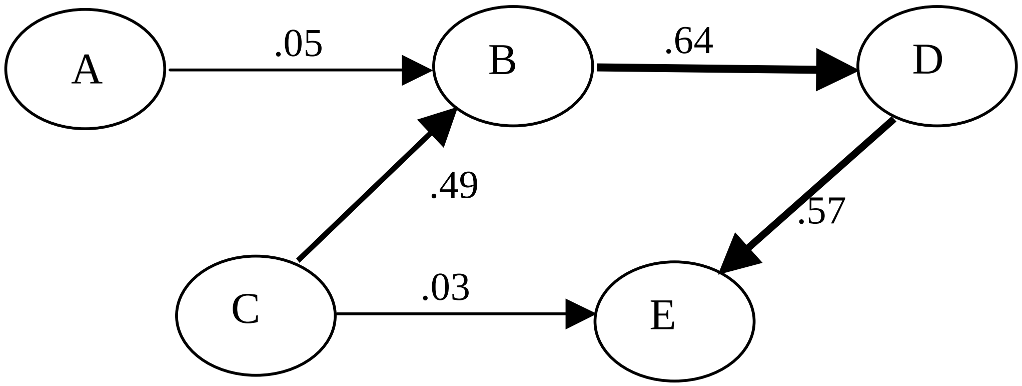
***Weighted Relevance* is symmetric for variables with the same dimension**

Representing expected relevance in influence diagrams

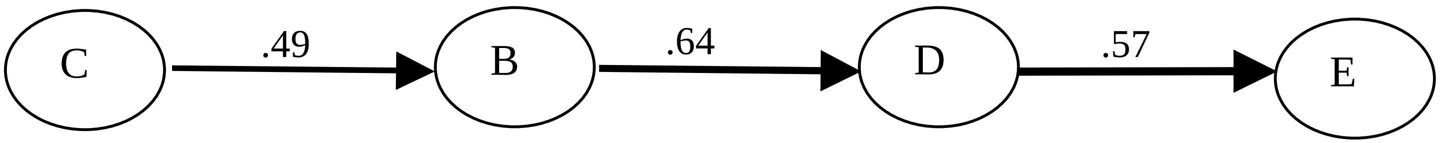


- **Thickness of relevance arrows reflect the amount of relevance (expected relevance value is written alongside the arrow).**
- **This measure is simple to compute and visualize.**

Weighted relevance as an aid in simplifying diagrams



This diagram can be simplified by dropping the low relevance links.



It is unlikely that low relevance links will make a large difference on the analysis.

In Short

1. Decision analysts have always had a feeling for the differences in relevance intensity, but no simple tool to quantify it.
2. Relevance can be measured by the distance between the prior probability distribution and the conditioned probability distributions.
3. Weighted Relevance gives us the overall importance of the probabilistic link.
4. These indicators can enrich the representation of complex business decisions, and improve communication between analysts and their clients.