PETRÓLEO BRASILEIRO S.A.
PETROBRAS

Pontifícia Universidade Católica do Rio de Janeiro

Real Options, Learning Measures, and Bernoulli Revelation Processes

9th Annual International Conference on Real Options
June 22-25, 2005 - Paris, France

By: Marco Antonio Guimarães Dias, Doctor
– Senior Consultant (internal) by Petrobras, Brazil
– Adjunct Professor of Finance (part-time) by PUC-Rio

Revised Paper copy? Send e-mail to marcoagd@pobox.com
Visit the first real options website: www.puc-rio.br/marco.ind/

Presentation Outline

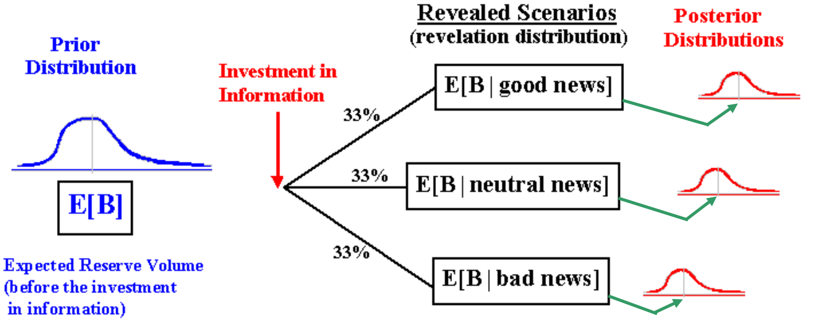
- Technical uncertainty modeling and learning options
- *Revelation distribution* theorem to solve complex real options problems with both technical and market uncertainties into a Monte Carlo framework.
- A theory on *learning measures* to ease the solution of problems with learning options.
  - A proposed learning measure η^2, the expected % of variance reduction = normalized variance of revelation distribution.
    - This learning measure has nice/adequate mathematical properties.
  - Axiom list for learning measures.
- Motivating example for portfolio of oil exploration assets.
- Bernoulli revelation processes: the simplest event-driven diffusion process for technical uncertainty evolution.
  - Used to model oil exploration portfolio; long-term valuation to invest in new petroleum basins; option games applications; R&D projects.

Technical Uncertainty and Learning Options

- **Technical uncertainty is related with specific project factors.**
  - **It incentives the *investment in learning processes* about the profit function: *value of information* (VOI) or *learning options* models.**
- **In petroleum E&P is the uncertainty related with the *existence* (chance factor), *volume* (B) and *quality* (q) of an oil/gas reserve.**
- **In corporations with diversified investors, <u>technical uncertainty does not demand risk-premium</u> because has zero correlation with the market portfolio (so, $\beta = 0$ in CAPM world).**
  - **But <u>technical uncertainty reduces the project NPV (and real option</u> value) due to <u>suboptimal investment</u> and *not* due to manager utility.**
  - **In terms of optimal investment scale, technical uncertainty almost surely will cause either *under*-investment or *over*-investment.**
- **There are some papers using stochastic processes like GBM to model technical uncertainty: this is very inadequate. Example:**
  - **Expectations about the volume B changes only when a *learning option* is exercised (new information), whereas oil prices change everyday.**

Learning Options and Changes in Expectations

- **In general, learning option exercise (investment in information) changes the original (prior) uncertainty and the expected value.**



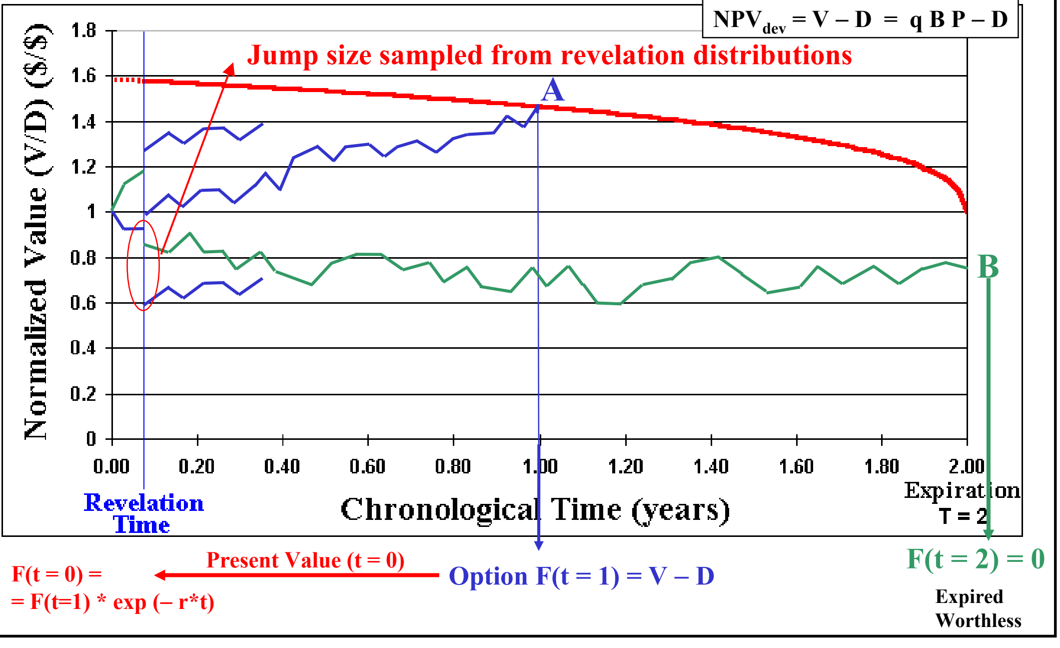
- **We´ll show that new relevant information *always* reduces the <u>average</u> uncertainty (expected variance of posterior distributions)**
- **We don´t know the true values from posterior distributions, so usually we work with the expected values to optimize the project.**
  - **So, it´s natural to work with the *distribution of conditional expectations*, here named *revelation distribution*. Here is developed a theory for this.**
  - **Note that the revelation distribution is unique even with infinite posterior distrib.**

The Revelation Distribution Properties

- ***Full Revelation*** **(perfect information): occurs when the signal S (new information) reveals *all the truth* on a variable of interest X.**
  - **Much more common is the *partial revelation* case, but full revelation is important as the limit goal for any investment in information process.**
- **Theorem 1: The revelation distribution of $R_X = E[X \mid S]$ have 4 nice properties (limit, mean, variance and martingale):**
  - **a)** **For the *limit* case of full revelation, the revelation distribution $p(R_X)$ is equal to p(x), the prior distribution of X. So, the R_X *variance is bounded*.**
  - **b)** **The revelation distribution expected value is equal to the original (prior) distribution expected value.**
    - $E[E[X \mid S]] = E[R_X] = E[X]$ **(known as *law of iterated expectations*)**
  - **c)** **The revelation distribution variance is equal to the expected variance reduction (i.e., prior variance less the expected posterior variance).**
    - $Var[E[X \mid S]] = Var[R_X] = Var[X] - E[Var[X \mid S]]$ **= Expected variance reduction (this property reports the *revelation power* of a learning project).**
  - **d)** **In a sequential learning process $S_1, S_2, \ldots$ the *revelation process* $\{R_{X,n}\} = \{E[X \mid S_1, S_2, \ldots S_n]\}$, n = 1, 2, … is a (event-driven) martingale.**
    - **In short, ex-ante these random variables $R_{X,n}$ have the same mean.**

Oilfield with Technical and Market Uncertainties

- **MC simulation combining uncertainties and using R_X distrib. (Dias, 2002).**

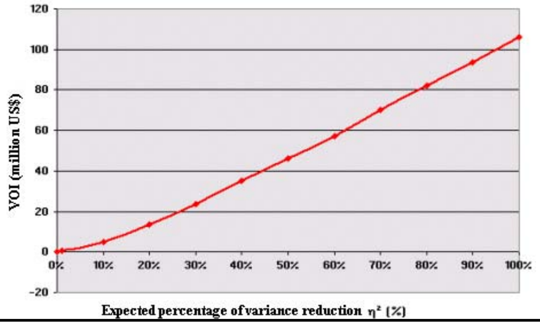


Proposed Learning Measure η^2

- The paper proposes a learning measure (η^2) linked with the revelation distribution variance given by Theorem 1(c).
  - The learning measure η^2 is the *expected percentage of variance reduction*;
  - η^2 is also the *normalized* revelation distribution variance.

$$\eta^2(X \mid S) = \frac{\mathrm{Var}[X] - E[\,\mathrm{Var}[X \mid S]\,]}{\mathrm{Var}[X]} = \frac{\mathrm{Var}[\,E[X \mid S]\,]}{\mathrm{Var}[X]} = \frac{\mathrm{Var}[R_X]}{\mathrm{Var}[X]}$$

- It has surprising nice properties (next slide) and will be used also to build revelation processes. (jump?)
- It can be used in Dias (2002) example to see the VOI x η^2:



- It is rough linear, except for nonconcavity at low η^2.
- Because is easier to optimize with perfect information, we could use linearity (VOI at $\eta^2 = 0$ and $\eta^2 = 1$) for fast calculus
- This was done at Petrobras for 4D seismic valuation (2003).

Properties for the Learning Measure η^2

- **Proposition 1**: Let X and S be two non-trivial r.v. with finite variances and learning measure $\eta^2(X \mid S)$ defined before. Then:
  a) The measure $\eta^2(X \mid S)$ always exists;
  b) The measure η^2 is, in general, asymmetric, i.e., $\eta^2(X \mid S) \neq \eta^2(S \mid X)$;
  c) The measure η^2 is normalized in unit interval, i.e., $0 \leq \eta^2 \leq 1$;
  d) If X and S are independent r.v. $\Rightarrow$ $\eta^2(X \mid S) = \eta^2(S \mid X) = 0$; in addition, is valid the condition: $\eta^2(X \mid S) = 0 \Leftrightarrow \mathrm{Var}[R_X(S)] = 0$;
  e) $\eta^2(X \mid S) = 1 \Leftrightarrow$ exists a functional dependence, i.e., $\exists$ r.v. g(S), so that X = g(S);
  f) The measure $\eta^2(X \mid S)$ é invariant under linear transformations of X, i.e., for any real numbers a and b, with $a \neq 0$, $\eta^2(a\,X + b \mid S) = \eta^2(X \mid S)$;
  g) $\eta^2(X \mid S)$ is invariant under linear and nonlinear transformation of S if the transformation g(S) is a 1-1 function: $\eta^2(X \mid g(S)) = \eta^2(X \mid S)$, if g(s) is invertible;
  h) If $Z_1, Z_2, \ldots$ are iid, with $S = Z_1 + \ldots + Z_j$ and $X = Z_1 + \ldots + Z_{j+k} \Rightarrow \eta^2(X \mid S) = j / (j+k)$
- **Theorem 2** (Learning measure decomposition): Let the signals $S_1, S_2, \ldots, S_n$, be *independent*. Let $X = f(S_1) + g(S_2) + \ldots + h(S_n)$, where f, g, . . . , h, are any real valued functions. Then:

$$\eta^2(X \mid S_1) + \eta^2(X \mid S_2) + \ldots + \eta^2(X \mid S_n) = 1$$

- For product or quotient of functions, we can apply it with a logarithm transform.

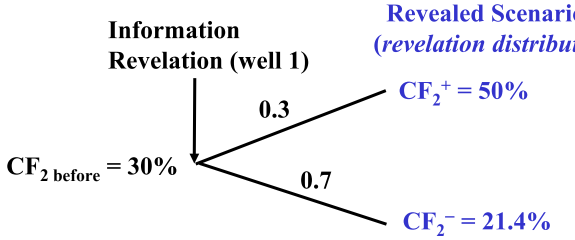
Axioms for Probabilistic Learning Measures

- Inspired in Rényi axioms for measures of dependence between r.v., the axiom list below gives the desirable properties for a learning measure $M(X \mid S)$:
  - A. $M(X \mid S)$ shall <u>exist</u> at least for all *non-trivial* r.v. X and S with *finite* uncertainties;
  - B. $M(X \mid S)$ shall be, in general, <u>asymmetric</u> in order to capture learning asymmetries between X and S (X can learn a lot with S, but not vice-versa, e.g., $X = S^2$);
  - C. $M(X \mid S)$ shall be <u>normalized</u> in the unit interval, i. e., $0 \leq M(X \mid S) \leq 1$;
  - D. $M(X \mid S) = 0 \Rightarrow$ <u>zero learning</u> (including, but not only, if X and S are independent);
  - E. In case of <u>functional dependence</u>, M shall be maximum: $X = f(S) \Rightarrow M(X \mid S) = 1$. In addition, $M(X \mid S) = 1 \Rightarrow$ <u>maximum learning</u>;
  - F. $M(X \mid S)$ shall be invariant under linear transforms (changes of scale) of either X or S, i.e. ($a \neq 0$): $M(a X + b \mid S) = M(X \mid S)$ and $M(X \mid S) = M(X \mid a S + b)$;
  - G. $M(X \mid S)$ shall be <u>practical</u>, i.e., *easy interpretation* and *easy to quantify/estimate*;
  - H. $M(X \mid S)$ shall be <u>additive</u>, i. e., if S can be decomposed into a sum of *independent* factors $S_1 + S_2 + \ldots + S_n$, with knowledge of all S_k providing <u>maximum learning</u>, then: $M(X \mid S_1) + M(X \mid S_2) + \ldots + M(X \mid S_n) = 1$
- <u>Theorem 3</u>: The proposed learning measure η^2 obeys the entire axiom list. In some cases (F and H), in a way stronger than required.
  - Showed with Prop. 1 and Theorem 2, considering the *variance as learning indicator*.

Bernoulli Revelation Process: Motivating Example

- One exploratory tract has two *positively correlated* and *equal* prospects, both with *chance factor* CF = 30%, drilling cost of \$ 30 million, and both with NPV_{DP} = 95 million, in case of success.
  - So, both the *expected monetary values* (EMV) are negatives:
    $EMV_1 = EMV_2 = -30 + (0.3 \times 95) = -1.5$ million
    - For simplicity, don´t consider market uncertainty (in oil prices, etc.)
  - Is the portfolio value = zero? No, the prospects are dependent: drilling first the prospect 1 reveals information to prospect 2!
    - Due to prospects positive dependence, CF_2 *must* be revised $CF_2^+ > CF_2$ in case of good news and down to $CF_2^- < CF_2$, in case of bad news.
    - Let the dependence of CF_1 (the signal) and CF_2 (interest variable) be so that:

Revealed Scenarios
(*revelation distribution*)



Updated EMV_2

$EMV_2^+ = -30 + (0.5 \times 95) = +17.5$ million

$EMV_2^- = -30 + (0.214 \times 95) = -9.7$ million

Dependence Leverages Value: Oil Exploration Portfolio

- Because we have an *option* to drill the well 2, we will exercise this option only if the revealed EMV_2 is positive (i.e. EMV_2^+)
  - So, the portfolio value Π is positive thanks to optionallity and dependence

$\Pi = EMV_1 + [CF_1 \,.\, Max(0, EMV_2^+)] + [(1 - CF_1) \,.\, Max(0, EMV_2^-)] \Rightarrow$

$\Rightarrow \Pi = -1.5 + [(30\% \text{ x } 17.5) + (70\% \text{ x } 0)] = +3.75$ MM$

  - A very different value when compared with the case of independence
- So, for *portfolio of real assets*, dependence can leverage value.
  - In addition, if one firm owns one prospect and other firm owns the other one, the revelation effect is also important for *option games applications* (war of attrition and bargain), as showed by Dias & Teixeira (2004).
  - A real life application at Petrobras (2003) was the long-term valuation of entry in a new petroleum basin, considering information revelation effects.
- Chance factor has a Bernoulli distribution: two scenarios (0 or 1), with probability of success (i.e., CF = 1) of p.
- The prospects dependence causes a *learning effect*, which must be studied in the context of *bivariate Bernoulli distributions*.
  - This study set the revealed scenarios CF^+ and CF^- as function of η^2

Bivariate Bernoulli Distribution Biv-Be(X, S)

- The bivariate Bernoulli distribution and its marginals are showed below:

		Signal S (eg.: seismic)		Marginal distribution of X (CF)
		S = 1	S = 0	
Variable X (eg.: chance factor)	X = 1	p_{11}	p_{10}	p
	X = 0	p_{01}	p_{00}	1 – p
Marginal Distribution of S		q	1 – q	

- Multivariate distribution literature shows that are necessary *limits of consistence* for these distributions, named Fréchet-Hoeffding bounds.
  - In particular is not possible *full revelation* for any marginals X, S.
- Without loss of generality, set X = CF and by notational convenience for Bernoulli revelation processes, set $p = CF_0$. In this case, the revelation distribution has two scenarios denoted by CF^+ and CF^-:

$$CF^+ = Pr[CF = 1 \mid S = 1] = E[CF \mid S = 1]$$
$$CF^- = Pr[CF = 1 \mid S = 0] = E[CF \mid S = 0]$$

  - So, CF_0 *evolves* to CF^+ or CF^-, depending on the signal S.
  - Elementary probability and η^2 definition are used in the following theorem.

Revealed CF Scenarios and Learning Measure η^2

- **Theorem 4**: Let the (non-trivial) marginals be $CF \sim Be(CF_0)$ and $S \sim Be(q)$, w/ a bivariate Bernoulli distribution linked by the learning measure $\eta^2(CF\,|\,S)$
  - The revealed success probabilities CF^+ and CF^-, for *positive* dependence, are:

$$CF^+ = CF_0 + \sqrt{\frac{1-q}{q}}\sqrt{CF_0\,(1-CF_0)}\,\sqrt{\eta^2(CF\,|\,S)}$$

$$CF^- = CF_0 - \sqrt{\frac{q}{1-q}}\sqrt{CF_0\,(1-CF_0)}\,\sqrt{\eta^2(CF\,|\,S)}$$

For negative dependence, just change signal after CF_0

  - η^2 in this case is equal to the square of correlation coefficient ρ:

$$\eta^2(CF\,|\,S) = \rho^2(CF, S) = \frac{(p_{11} - CF_0\,q)^2}{CF_0\,(1-CF_0)\,q\,(1-q)}$$

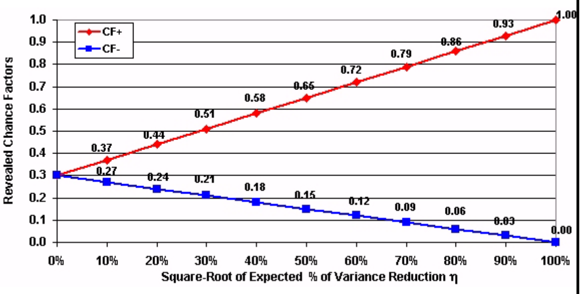
  - Here η^2 is symmetric: X and S ~ Bernoulli $\Rightarrow \eta^2(CF\,|\,S) = \eta^2(S\,|\,CF)$
  - Here (Bernoulli) $\eta^2(CF\,|\,S) = 0 \Leftrightarrow$ CF and S independent
  - To assure the bivariate Bernoulli distrib. existence, the Fréchet-Hoeffding bounds in terms of η^2 are:

$$0 \le \eta^2(CF\,|\,S) \le \mathrm{Max}\left\{ \mathrm{Max}\left\{\frac{CF_0\,q}{(1-CF_0)\,(1-q)}, \frac{(1-CF_0)\,(1-q)}{CF_0\,q}\right\}, \frac{\mathrm{Min}\{CF_0, q\}\,(1-\mathrm{Max}\{CF_0, q\})}{\mathrm{Max}\{CF_0, q\}\,(1-\mathrm{Min}\{CF_0, q\})} \right\}$$

Exchangeable Bernoulli Distributions

- An important simplification is when the r.v. $CF \sim Be(CF_0)$ and $S \sim Be(q)$ are *exchangeable* (here mean $p_{01} = p_{10}$). This assumption has been used intuitively in professional petroleum literature (e.g. Wang et al, 2000).
- **Proposition 2**: if CF and S are exchangeable, then:
  - CF and S exchageable $\Leftrightarrow CF_0 = q$
  - The Fréchet-Hoeffding bounds are not restrictions anymore: $0 \le \eta^2 \le 1$
  - The revealed success probabilities CF^+ and CF^-, in case of *positive* dependence, are:

$$CF^+ = CF_0 + (1 - CF_0)\,\eta$$
$$CF^- = CF_0 - CF_0\,\eta$$
$$\Rightarrow CF^+ - CF^- = \eta$$



- **Lemma 3**: The necessary condition for CF full revelation (maximum learning) is that CF and S be exchangeable r.v.:

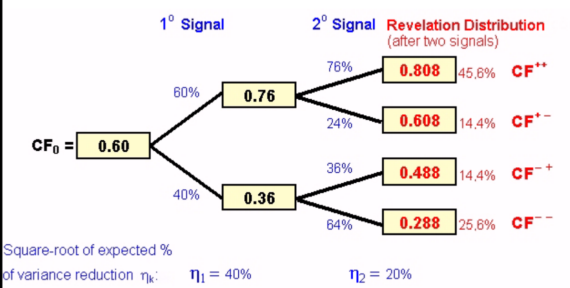
$\eta^2(CF\,|\,S) = 1 \Rightarrow$ CF and S exchangeable r.v. (jump next?)

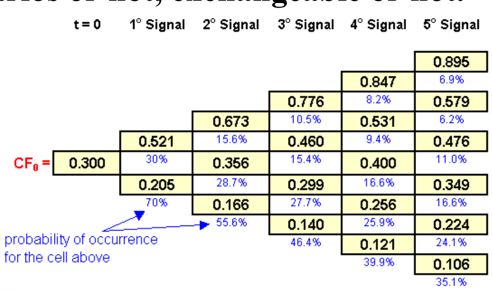
Bernoulli Revelation Processes: Definitions

- <u>Exploratory discovery process</u>: is a sequence of *learning options exercises* that results in a discovery. In general, these learning options have different *learning costs*, different *time to learn*, and different *revelation powers,* that can measured by η^2.
  - In oil exploration this sequence can be seismic search, drilling of wildcat wells, etc. In R&D, this sequence can be the R&D phases.
- <u>Exploratory revelation process</u>: is the probabilistic effect over the interest variable caused by the exploratory discovery process. This variable can be the chance factor, oil in place, etc.
- <u>Bernoulli revelation process</u>: is a *sequence of revelation distributions generated by a sequence of bivariate Bernoulli distributions* that represents the interaction between a sequence of signals S with the chance factor of interest CF.
  - In particular, *exchangeable* revelation Bernoulli processes permits that, given CF_0 and a sequence of η_k, k = 1, 2, …, we can construct all the revelation process because the success probabilities for the signals S_k are automatically defined by exchangeability assumption.

Bernoulli Revelation Processes: Types

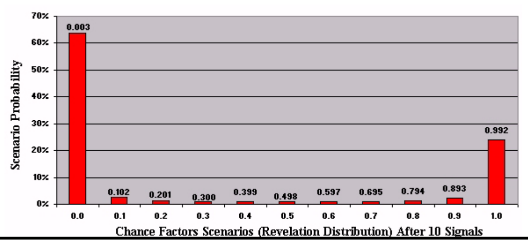
- Bernoulli Revelation Processes can be totally convergente (to full revelation) or not; recombining scenarios or not, exchangeable or not.





- An apparent paradox: By Theorem 1(a), the revelation distribution at full revelation limit must have only two scenarios (prior distrib.) How is it true if is growing the nº of scenarios?

Near of full revelation limit, the mass probability concentrates at scenarios near 0 and near 1. In limit, only two scenarios has mass probability, and = prior prob.



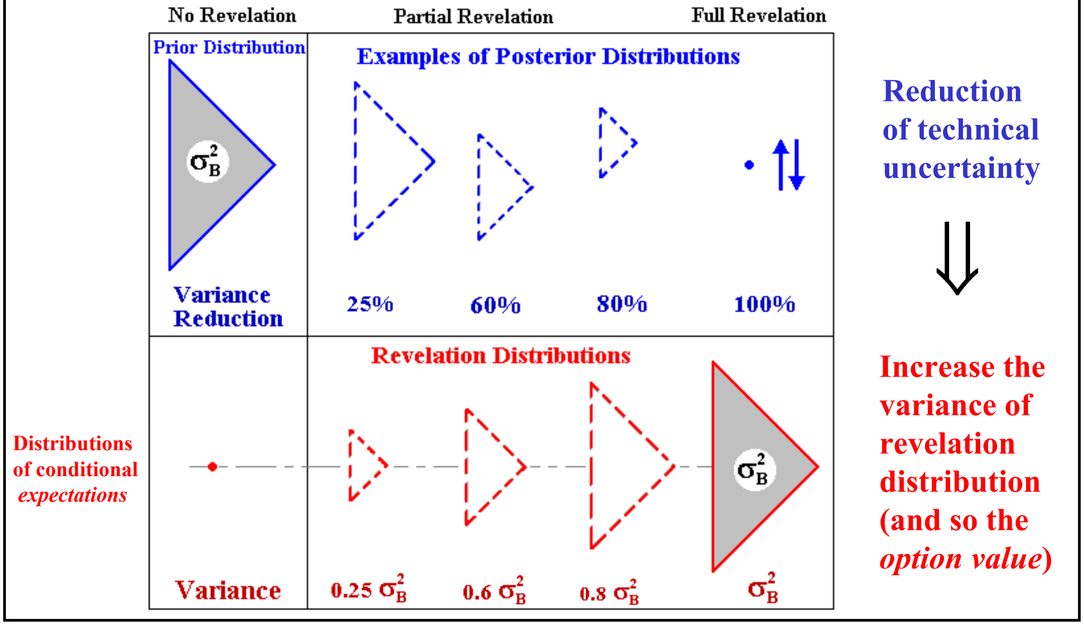
Conclusions

- **This paper presents a theory for technical uncertainty and its evolution with the exercise of learning options, based in the conditional expectations generated by information revelation.**
  - **Theorem 1 pointed the main conditional expectation distribution properties. This distribution is named revelation distribution.**
    - **Revelation distributions was used to solve complex real options problems in a risk-neutral Monte Carlo framework combining many uncertainties.**
- **The proposed learning measure η^2 has a direct link with that because is the normalized revelation distribution variance.**
  - **It has the intuitive interpretation as the expected % of variance reduction induced by the exercise of learning options.**
  - **Surprising convenient math properties were presented for η^2.**
- **This theory was applied to oil exploration chance factor, starting a theory on Bernoulli revelation processes (BRP).**
  - **After a motivating example in portfolio of exploratory prospects, was discussed properties of BRP and some possible alternatives.**
- **Thank you very much! Merci beaucoup!**

APPENDIX
SUPPORT SLIDES

Posterior Distributions x Revelation Distribution

- Higher volatility, higher option value. Why learn, reducing uncertainty?

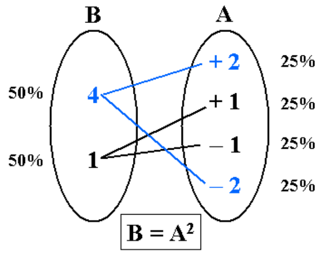


Conditional Expectation in Theory and Practice

- What is the relevance of the *conditional expectation* concept for *learning* process *valuation*?
  - We saw that is much simpler to work with the *unique* conditional expectation distribution than *many* posterior distributions.
  - Other practical advantage: expectations has a natural place in finance
    - Firms use current expectations to calculate the NPV or the real options exercise payoff. Ex-ante the investment in information, the new *expectation is conditional.*
    - *The price of a derivative is simply an expectation of futures values* (Tavella, 2002)
  - The concept of conditional expectation is also *theoretically* sound:
    - We want to estimate X by observing I, using a function g(I).
    - The most frequent measure of quality of a predictor g is its *mean square error* defined by MSE(g) = $E[X - g(I)]^2$. The choice of g* that minimizes the error measure MSE(g) is exactly the conditional expectation E[X | I].
    - This is a very known property used in econometrics
  - Even in *decision analysis literature*, is common to work with conditional expectation inside the maximization equation (e.g., McCardle, 1985)
    - But instead *conditional expectation properties*, the focus has been *likelihood function*

Learning Asymmetry and Correlation

- Given two r.v. A and B, many times B learn more with A than A learn with B (or vice-versa), i.e., in general the learning is *asymmetric*. Let us see an example: $B = A^2$ for the numbers:



  - If we know the value of A, then the value of B is defined (full learning or full revelation).
  - But if we know the value of B, the value of A is not defined. (variance can even become higher).
- The correlation coefficient (symmetric measure) is zero (!) in the above example, although A and B be obviously dependent (they have a functional dependence!).
- The proposed measure η^2, is asymmetric and captures the full learning of B given A: $\eta^2(B \mid A) = 100\% \neq \eta^2(A \mid B) = 0\%$.

Information Likelihood

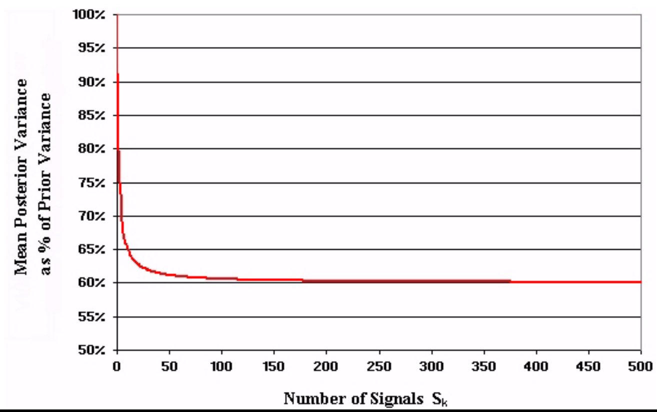
- In the traditional literature is much common the concept of information *likelihood* of S to predict X. It´s a kind of *reliability*.
  - They are the called *inverse probabilities* $p(S \mid X)$.
- However, likelihood is <u>not</u> a good learning measure. Example:
  - There are two "infallible experts" that know all the truth about the stock exchange market. They can be consulted in order to know if the company X stock is going to rise or not in the next day market. Assume that is known that one expert always says the truth and the other expert always lies: the inverse probabilities are $p_1(S = a \mid x = a) = 100\%$ and $p_2(S = a \mid x = a) = 0\%$, a = rise or not rise.
  - By the learning point of view the advises are equivalent, because we have full learning of X in both cases:
    - If the "infallible expert 2" says that the stock will not rise, then we learn that this stock is going to rise next day with probability 1!
  - So, this measure atributes <u>two different values for the same learning</u> and (even worse) using the <u>value zero to indicate maximum learning</u>.
  - Likelihood also doesn´t obey other elementar properties for M(X|S)

Recombining Process: Revelation Limit

- **Exchangeable recombining Bernoulli processes never converges to full revelation. In the limit, the process converges to a partial revelation, with the limit being simply $1 - \eta_1$, as showed by the limite:**

$$\lim_{n \to \infty} \prod_{k=1}^{n} \left(1 - \frac{\eta_1^2}{[1 + (k-1)\,\eta_1]^2} \right) = 1 - \eta_1$$

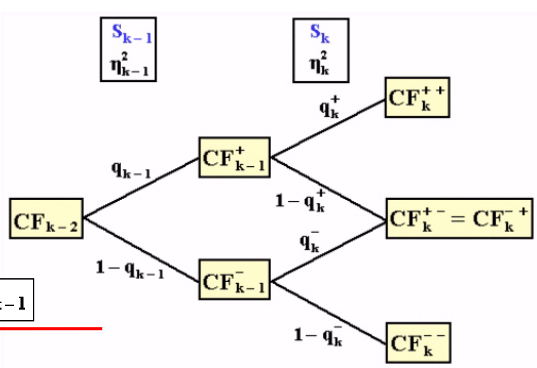
The figure shows an example with $\eta_1 = 0.4$



Bernoulli Revelation Processes: Some Details

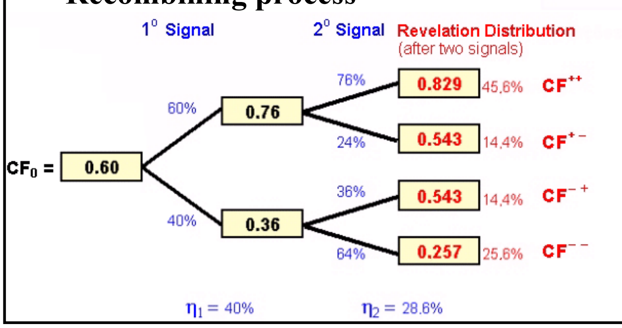
$$CF_k^{+-} = CF_{k-1}^{+} - \sqrt{\frac{q_k^+}{1-q_k^+}}\sqrt{CF_{k-1}^{+}\,(1-CF_{k-1}^{+})}\,\sqrt{\eta_k^2}$$

$$CF_k^{-+} = CF_{k-1}^{-} + \sqrt{\frac{1-q_k^-}{q_k^-}}\sqrt{CF_{k-1}^{-}\,(1-CF_{k-1}^{-})}\,\sqrt{\eta_k^2}$$



$$p_{i,k} = p_{i-1,k-1}\,(1 - CF_{i-1,k-1}) + p_{i,k-1}\,CF_{i,k-1}$$

Recombining process



Recombining, exchangeable: For signal S_k, we need only η_1 to get η_k:

$$\eta_k = \frac{\eta_1}{1 + (k-1)\,\eta_1}$$