

The Game-Theoretic Framework for Probability

Glenn Shafer

*Rutgers Business School — Newark and New Brunswick,
180 University Avenue, Newark, New Jersey 07102 USA*

*Department of Computer Science, Royal Holloway,
University of London, Egham, Surrey TW20 0EX UK
gshafer@rutgers.edu*

Abstract

The game-theoretic framework for probability was introduced by Vladimir Vovk and myself in *Probability and Finance*, published in 2001. But it has roots in the work of Jean Ville, who showed how Cournot's principle, the principle that an event of very small probability will not happen, can be given a game-theoretic interpretation. The framework allows us to generalize the classical limit theorems of probability and opens up new ways of making probability predictions.

1. Introduction

Cournot's principle says that an event of small or zero probability singled out in advance will not happen. From the turn of the twentieth century through the 1950s, many mathematicians, including Aleksandr Chuprov, Émile Borel, Maurice Fréchet, Paul Lévy, and Andrei Kolmogorov, saw this principle as fundamental to the application and meaning of probability. In their view, a probability model gains empirical content only when it rules out an event by assigning it small or zero probability.

In his doctoral dissertation, published in 1939, Jean Ville showed that Cournot's principle can be given a game-theoretic interpretation. The game-theoretic framework for probability published by Vladimir Vovk and myself in 2001 (in *Probability and Finance; It's Only a Game!*²⁰) extends Ville's game-theoretic approach to cases where successive forecasts fall short of a full probability distribution for the quantities forecast.

2. The Origins of Cournot's Principle

An event with very small probability is *morally impossible*; it will not happen. Equivalently, an event with very high probability is *morally certain*; it will happen. This principle was first formulated within mathematical probability by Jacob Bernoulli. In his *Ars Conjectandi*, published posthumously in 1713, Bernoulli proved that in a sufficiently long sequence of independent trials of an event, there is a very high probability that the frequency with which the event happens will be close to its probability. Bernoulli explained that we can treat the very high probability as moral certainty and so use the frequency of the event as an estimate of its probability.

Augustin Cournot, a mathematician now remembered as an economist and a philosopher of science,^{15,16} gave the discussion a nineteenth-century cast in his 1843 treatise on probability.⁹ Because he was familiar with geometric probability, Cournot could talk about probabilities that are vanishingly small. He brought physics to the foreground. It may be mathematically possible, he argued, for a heavy cone to stand in equilibrium on its vertex, but it is physically impossible. The event's probability is vanishingly small. Similarly, it is physically impossible for the frequency of an event in a long sequence of trials to differ substantially from the event's probability [9, pp. 57 and 106].

At the turn of the twentieth century, it was a commonplace among statisticians that one must decide what level of probability will count as practical certainty in order to apply probability theory. We find this stated explicitly in 1901, for example, in the articles by Georg Bohlmann and Ladislaus von Bortkiewicz in the section on probability in the *Encyklopädie der mathematischen Wissenschaften* [25, p. 825] [4, p. 861]. Aleksandr Chuprov, professor of statistics in Petersburg, was the champion of Cournot's principle in Russia.

He called it Cournot's lemma [18, p. 167] and declared it a basic principle of the logic of the probable [22, pp. 95–96].

Saying that an event of very small or vanishingly small probability will not happen is one thing. Saying that probability theory gains empirical meaning only by ruling out the happening of such events is another. Cournot may have been the first to make this second assertion:

... The physically impossible event is therefore the one that has infinitely small probability, and only this remark gives substance — objective and phenomenal value — to the theory of mathematical probability [9, p. 78].

Paul Lévy, a French mathematician who began writing on probability in the 1920s, stands out for the clarity of his articulation of the thesis that Cournot's principle is the only way of connecting a probabilistic theory with the world outside mathematics.¹⁴ Lévy's views were widely shared in France. In the 1940s, Émile Borel called Cournot's principle first "the only law of chance" (la loi unique du hasard).^{5,6} Neither Lévy nor Borel used the name "Cournot's principle," which was coined by Maurice Fréchet in 1949. Fréchet's inspiration was Oskar Anderson, who had talked about the Cournotsche Lemma (Cournot's lemma) and the Cournotsche Brücke (Cournot's bridge).^{1,2} Anderson was following his teacher Chuprov in the use of "lemma." Fréchet felt that "lemma," like "theorem," should be reserved for purely mathematical results and so suggested "principe de Cournot." Fréchet's coinage was used in the 1950s in French, German, and English.^{12,18,19,26}

3. Ville's Theorem

Vovk and I²⁰ use Cournot's principle in a game-theoretic form: a strategy for placing bets without risking bankruptcy will not multiply the bettor's capital by a large or infinite factor. In the case where the bettor can buy or sell any random variable for its expected value, this is equivalent to the classical form of the principle; Jean Ville demonstrated the equivalence in 1939.²⁴

Consider a sequence Y_1, Y_2, \dots of binary random variables with a joint probability distribution P . Suppose, for simplicity, that P

assigns every finite sequence y_1, \dots, y_n of 0s and 1s positive probability, so that its conditional probabilities for Y_n given values of the preceding variables are always unambiguously defined. Following Jean Ville,²⁴ consider a gambler who begins with \$1 and is allowed to bet as he pleases on each round, provided that he does not risk bankruptcy. We can formalize this with the following protocol, where betting on Y_n is represented as buying some number s_n (possibly zero or negative) of tickets that cost $\$P\{Y_n = 1 | Y_1 = y_1, \dots, Y_{n-1} = y_{n-1}\}$ and pay $\$Y_n$.

BINARY PROBABILITY PROTOCOL

Players: Reality, Skeptic

Protocol:

$$\mathcal{K}_0 := 1.$$

FOR $n = 1, 2, \dots$:

Skeptic announces $s_n \in \mathbb{R}$.

Reality announces $y_n \in \{0, 1\}$.

$$\mathcal{K}_n := \mathcal{K}_{n-1} + s_n(y_n - P\{Y_n = 1 | Y_1 = y_1, \dots, Y_{n-1} = y_{n-1}\}).$$

Restriction on Skeptic: Skeptic must choose the s_n so that his capital is always non-negative ($\mathcal{K}_n \geq 0$ for all n) no matter how Reality moves.

This is a perfect-information sequential protocol; moves are made in the order listed, and each player sees the other player's moves as they are made. The sequence $\mathcal{K}_0, \mathcal{K}_1, \dots$ is Skeptic's capital process.

Ville showed that Skeptic's getting rich in this protocol is equivalent to an event of small probability happening, in the following sense:

- (1) When Skeptic follows a measurable strategy (a rule that gives s_n as a function of y_1, \dots, y_{n-1}),

$$P\left\{\sup_n \mathcal{K}_n \geq \frac{1}{\epsilon}\right\} \leq \epsilon \quad (1)$$

for every $\epsilon > 0$. (This is because the capital process $\mathcal{K}_0, \mathcal{K}_1, \dots$ is a non-negative martingale; Equation (1) is sometimes called *Doob's inequality*.)

- (2) If A is a measurable subset of $\{0, 1\}^\infty$ with $P(A) \leq \epsilon$, then Skeptic has a measurable strategy that guarantees

$$\liminf_{n \rightarrow \infty} \mathcal{K}_n \geq \frac{1}{\epsilon}$$

whenever $(y_1, y_2, \dots) \in A$.

We can summarize these results by saying that Skeptic's being able to multiply his capital by a factor of $1/\epsilon$ or more is equivalent to the happening of an event with probability ϵ or less.

Ville's work was motivated by von Mises's notion of a collective.^{27–29} Von Mises had argued that a sequence y_1, y_2, \dots of 0s and 1s should be considered random if no subsequence with a different frequency of 1s can be picked out by a gambler to whom the y s are presented sequentially; this condition, von Mises felt, would keep the gambler from getting rich by deciding when to bet. Ville showed that von Mises's condition is insufficient, in as much as it does not rule out the gambler's getting rich by varying the direction and amount to bet.

4. The Game-Theoretic Framework

Although the preceding explanation of Ville's ideas was limited to the binary case, Ville made it clear that these ideas apply whenever conditional probabilities from a joint probability distribution for a sequence of random variables are used to make successive probability predictions. The framework of²⁰ generalizes the ideas further. The generalization has three aspects:

- Instead of beginning with a probability measure and using its conditional probabilities or expected values as prices on each round, we allow another player, Forecaster, to set the prices as play proceeds. This makes the framework “prequential”;¹¹ there is no need to specify what the price on the n th round would be had Reality moved differently on earlier rounds.
- When convenient, we make explicit additional information, say x_n , that Reality provides to Forecaster and Skeptic before they make their n th moves.
- We allow the story to be multi-dimensional, with Reality making several moves and Forecaster pricing them all.

A convenient level of generality for the present discussion is provided by the following protocol, where \mathbb{R}^k is k -dimensional Euclidean space, \mathbf{Y} is a subset of \mathbb{R}^k , and \mathbf{X} is an arbitrary set.

LINEAR FORECASTING PROTOCOL

Players: Reality, Forecaster, Skeptic

Protocol:

$$\mathcal{K}_0 := 1.$$

FOR $n = 1, 2, \dots, N$:

Reality announces $x_n \in \mathbf{X}$.

Forecaster announces $f_n \in \mathbb{R}^k$.

Skeptic announces $s_n \in \mathbb{R}^k$.

Reality announces $y_n \in \mathbf{Y}$.

$$\mathcal{K}_n := \mathcal{K}_{n-1} + s_n \cdot (y_n - f_n).$$

Restriction on Skeptic: Skeptic must choose the s_n so that his capital is always non-negative ($\mathcal{K}_n \geq 0$ for all n) no matter how the other players move.

Here $s_n \cdot (y_n - f_n)$ is the dot product of the k -dimensional vectors s_n and $y_n - f_n$. Notice also that play stops on the N th round rather than continuing indefinitely. This is a convenient assumption in this section, where we emphasize the finitary picture; we will return to the infinitary picture later.

The linear forecasting protocol covers many prediction problems considered in statistics (where x and y are often called *independent* and *dependent* variables, respectively) and machine learning (where x is called the *object* and y the *label*).^{13,23,31} Market games can be included by taking f_n to be a vector of opening prices and y_n the corresponding vector of closing prices for the n th trading period.

A strategy for Skeptic in the linear forecasting protocol is a rule that gives each of his moves s_n as a function of the preceding moves by Reality and Forecaster, $(x_1, f_1, y_1), \dots, (x_{n-1}, f_{n-1}, y_{n-1}), x_n, f_n$. A strategy for Forecaster is a rule that gives each of his moves f_n as a function of the preceding moves by Reality and Skeptic, $(x_1, s_1, y_1), \dots, (x_{n-1}, s_{n-1}, y_{n-1}), x_n$. One way of prescribing a strategy for Forecaster is to choose a probability distribution for $(x_1, y_1), (x_2, y_2), \dots$ and set f_n equal to the conditional expected value of y_n given $(x_1, y_1), \dots, (x_{n-1}, y_{n-1}), x_n$. We will look at other interesting strategies for Forecaster in Sec. 7.

How can one express confidence in Forecaster? The natural way is to assert Cournot's principle: say that a legal strategy for Skeptic (one that avoids $\mathcal{K}_n < 0$ no matter how the other players move) will not multiply Skeptic's initial capital by a large factor.

Once we adopt Cournot's principle in this form, it is natural to scale the implications of our confidence in Forecaster the same way we do in classical probability. This means treating an event that happens only when a specified legal strategy multiplies the capital by $1/\epsilon$ as no more likely than an event with probability ϵ .

As in classical probability, we can combine Cournot's principle with a form of Bernoulli's theorem to obtain a statement about relative frequency in a long sequence of events. In a sufficiently long sequence of events with upper probability 0.1 or less, for example, it is morally certain that no more than about 10% of the events will happen [21, Sec. 5.3]. This is a martingale-type result; rather than insist that the events be independent in some sense, we assume that the upper probability for each event is calculated at the point in the game where the previous event is settled.

5. Extending the Classical Limit Theorems

One of the main contributions of Shafer and Vovk²⁰ was to show that game theory can replace measure theory as a foundation for classical probability.

We showed in particular that classical limit theorems, especially the strong law of large numbers and the law of the iterated logarithm, can be proven constructively within a purely game-theoretic framework. From Ville's work, we know that for any event with probability zero, there is a strategy for Skeptic that avoids bankruptcy for sure and makes him infinitely rich if the event fails. But constructing the strategy is another matter. In the case of the events of probability zero associated with the classical theorems, we did construct the requisite strategies; they are computable and continuous.

We provided similar constructions for classical results that do not require an infinite number of rounds of play to be meaningful: the weak law of large numbers, finitary versions of the law of the iterated logarithm, and the central limit theorem.

6. The Idea of a Quasi-Universal Test

If two events have very small probability, their union also has reasonably small probability. The analogous idea in game-theoretic probability is that of averaging strategies: if one strategy for Skeptic makes him very rich without risking bankruptcy if one event happens, and another makes him very rich without risking bankruptcy if a second event happens, then the average of the two strategies will make reasonably rich without risking bankruptcy if either of the events happens. This leads us to the notion of a quasi-universal strategy: we list the most important extreme events that we want to rule out, and by averaging the strategies that rule each out, we obtain a strategy that rules them all out.

Leaving aside how this idea has been developed in the past within measure-theoretic probability, let us consider how it can be developed measure-theoretically in this protocol of binary forecasting:

BINARY PROBABILITY PROTOCOL WITH FORECASTER AND OBJECTS

Players: Reality, Forecaster, Skeptic

Protocol:

$$\mathcal{K}_0 := 1.$$

FOR $n = 1, 2, \dots$:

Reality announces $x_n \in \mathbf{X}$.

Forecaster announces $p_n \in [0, 1]$.

Skeptic announces $s_n \in \mathbb{R}$.

Reality announces $y_n \in \{0, 1\}$.

$$\mathcal{K}_n := \mathcal{K}_{n-1} + s_n(y_n - p_n).$$

Restriction on Skeptic: Skeptic must choose the s_n so that his capital is always non-negative ($\mathcal{K}_n \geq 0$ for all n) no matter how the other players move.

In this protocol, where Forecaster gives a probability p_n on each round, taking into account the previous outcomes y_1, \dots, y_{n-1} and auxiliary information x_1, \dots, x_n , we are mainly interested in two aspects of the agreement between the probabilities p_n and the outcomes y_n :

Calibration. Whenever there are a large number of rounds on which p_n is close to some fixed probability p^* , we want the frequency with which $y_n = 1$ on those rounds to be approximately equal to p^* .

Resolution. We want this approximate equality between frequency and p^* to remain true when we consider only rounds where p_n is close to p^* and also x_n is close to some fixed value x^* in the object space \mathbf{X} .

As it turns out,³⁴ we can often average strategies that reject Forecaster's performance over a grid of values of (x^*, p^*) that are sufficiently dense to capture all deviations of practical interest. This average strategy, which is testing for calibration and resolution, will not necessarily test for more subtle deviations by y_1, y_2, \dots from the forecasts p_1, p_2, \dots , such as those associated with the law of the iterated logarithm or Ville's refutation of von Mises's theory, but these more subtle deviations may hold little interest. So the average strategy can be regarded, for practical purposes, as a universal test. To avoid confusion, I call it a *quasi-universal strategy*.

7. Defensive Forecasting

In cases where we have a quasi-universal strategy, a new opportunity opens up for Forecaster. Forecaster will do well enough if he can avoid rejection by that strategy. Formally, he needs a winning strategy in a version of the game where Skeptic is required to follow the quasi-universal strategy but Reality is free to move as she pleases. Does Forecaster have such a winning strategy? The surprising answer is yes.

This is easiest to see in the case where the quasi-universal strategy gives a move for the n th round that is continuous in the forecast p_n . As it happens, this is not an unreasonable requirement. We can construct quasi-universal strategies for calibration and resolution that are continuous in this respect, and there is even a philosophical argument for ruling out any discontinuous strategy for Skeptic: discontinuous functions are not really computable.^{7,17}

As it turns out, it is easy to show that for any forecast-continuous strategy for Skeptic there exists a strategy for Forecaster that does not allow Skeptic's capital to grow, regardless of what Reality does. Let me repeat the simple proof given in Refs. 32 and 34. It begins by simplifying so that Forecaster's job seems to be even a little harder.

Instead of requiring that the entire forecast-continuous strategy for Skeptic be announced at the beginning of the game, we ask only that Skeptic announce his strategy for each round before Forecaster's move on that round. And we drop the restriction that Skeptic avoid risk of bankruptcy. This produces the following protocol:

BINARY FORECASTING AGAINST CONTINUOUS TESTS

Players: Reality, Forecaster, Skeptic

Protocol:

$$\mathcal{K}_0 := 1.$$

FOR $n = 1, 2, \dots$:

Reality announces $x_n \in \mathbf{X}$.

Skeptic announces continuous $S_n : [0, 1] \rightarrow \mathbb{R}$.

Forecaster announces $p_n \in [0, 1]$.

Reality announces $y_n \in \{0, 1\}$.

$$\mathcal{K}_n := \mathcal{K}_{n-1} + S_n(p_n)(y_n - p_n).$$

Here S_n is Skeptic's strategy for the n th round; it gives his move as a function of Forecaster's not-yet-announced move p_n .

Theorem 1. *Forecaster has a strategy that ensures $\mathcal{K}_0 \geq \mathcal{K}_1 \geq \mathcal{K}_2 \geq \dots$.*

Proof Because S_n is continuous, Forecaster can use the following strategy:

- if the function $S_n(p)$ takes the value 0, choose p_n so that $S_n(p_n) = 0$;
- if S_n is always positive, take $p_n := 1$;
- if S_n is always negative, take $p_n := 0$.

This guarantees that $S_n(p_n)(y_n - p_n) \leq 0$, so that $\mathcal{K}_n \leq \mathcal{K}_{n-1}$. \square

Some readers may question the philosophical rationale for requiring that S_n be continuous. As it turns out, dropping this requirement does not cost us much; Forecaster can still win if we allow him to randomize.³³ This means that instead of telling Reality his probability p_n , Forecaster may give Reality only a probability distribution P_n for p_n , with the value p_n to be drawn from P_n out of sight of Reality or perhaps after Reality has selected y_n .

A strategy for Forecaster is what one usually calls a probability model; given the previous outcomes y_1, \dots, y_{n-1} and auxiliary

information x_1, \dots, x_n , it gives a probability p_n for $y_n = 1$. Such probabilities can be used in any repetitive decision problem.³⁰ So Theorem 1's guarantee that they are valid, in the sense that they pass any reasonable test of calibration and resolution, has immense practical significance.

References

1. O. N. Anderson, *Einführung in die mathematische Statistik*. Springer, Vienna (1935).
2. O. N. Anderson, Die Begründung des Gesetzes der grossen Zahlen und die Umkehrung des Theorems von Bernoulli. *Dialectica*, **3**(9/10), 65–77 (1949).
3. J. Bernoulli, *Ars Conjectandi*. Thurnisius, Basel, 1713. Edith Sylla's English translation, *The Art of Conjecturing, Together with Letter to a Friend on Sets in Court Tennis*, was published by Johns Hopkins University Press in 2005. Oscar Sheynin's English translation of Part IV, dated 2005, is at www.sheynin.de.
4. G. Bohlmann, Lebensversicherungs-Mathematik. In *Encyklopädie der mathematischen Wissenschaften, Bd. I, Teil 2*, pages 852–917. Teubner, Leipzig (1901).
5. É. Borel, *Les probabilités et la vie*. Presses Universitaires de France, Paris, 1943. Later editions in 1946, 1950, 1958, and 1967. The 1958 edition was translated into English by Maurice Baudin and published as *Probabilities and Life* by Dover, New York (1962).
6. É. Borel, *Probabilité et certitude*. Presses Universitaires de France, Paris, 1950. An English translation, *Probability and Certainty*, was published in 1963 by Walker, New York.
7. L. E. J. Brouwer, Begründung der Mengenlehre unabhängig vom logischen Satz vom ausgeschlossenen Dritte. Erster Teil. Allgemeine Mengelehre. *Koninklijke Nederlandse Akademie van Wetenschappen Verhandelingen*, **5**, 1–43 (1918).
8. A. A. Chuprov, Очерки по теории статистики (*Essays on the theory of statistics*). Sabashnikov, Saint Petersburg, second edition, 1910. The first edition appeared in 1909. The second edition was reprinted by the State Publishing House, Moscow (1959).
9. A.-A. Cournot, *Exposition de la théorie des chances et des probabilités*. Hachette, Paris, 1843. Reprinted in 1984 as Vol. I (B. Bru, ed.) of Ref. 1,
10. A.-A. Cournot, *Euvres complètes*. Vrin, Paris, 1973–1984. Vols. 10, with an eleventh to appear.
11. A. P. Dawid, Statistical theory: The prequential approach (with discussion). *Journal of the Royal Statistical Society. Series A*, **147**, 278–292 (1984).
12. B. de Finetti, Recent suggestions for the reconciliation of theories of probability. In Jerzy Neyman, editor, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 217–225. University of California Press, Berkeley and Los Angeles (1951).

13. T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning Theory: Data Mining, Inference, and Prediction*. Springer, New York (2001).
14. P. Lévy, *Calcul des probabilités*. Gauthier-Villars, Paris (1925).
15. T. Martin, *Probabilités et critique philosophique selon Cournot*. Vrin, Paris (1996).
16. T. Martin, *Bibliographie cournotienne*. Annales littéraires de l'Université Franche-Comté, Besançon (1998).
17. P. Martin-Löf, *Notes on Constructive Mathematics*. Almqvist & Wiksell, Stockholm (1970).
18. H. Richter, Zur Begründung der Wahrscheinlichkeitsrechnung. *Dialectica*, **8**, 48–77 (1954).
19. H. Richter, *Wahrscheinlichkeitstheorie*. Springer, Berlin (1956).
20. G. Shafer and V. Vovk, *Probability and Finance: It's Only a Game*. Wiley, New York (2001).
21. G. Shafer and V. Vovk, On a review by V. N. Tutubalin. http://www.probabilityandfinance.com/reviews/tut_response1.pdf (2002).
22. O. Sheynin, *Aleksandr A. Chuprov: Life, Work, Correspondence. The making of mathematical statistics*. Vandenhoeck & Ruprecht, Göttingen (1996).
23. V. Vapnik, *The Nature of Statistical Learning Theory*. Springer, New York (1996).
24. J.-A. Ville, *Étude critique de la notion de collectif*. Gauthier-Villars, Paris, 1939. This differs from Ville's dissertation, which was defended in March 1939, only in that a 17-page introductory chapter replaces the dissertation's one-page introduction.
25. L. von Bortkiewicz, Anwendungen der Wahrscheinlichkeitsrechnung auf Statistik. In *Encyclopädie der mathematischen Wissenschaften, Bd. I, Teil 2*, pages 821–851. Teubner, Leipzig (1901).
26. G. von Hirsch, Sur un aspect paradoxal de la théorie des probabilités. *Dialectica*, **8**, 125–144 (1954).
27. R. von Mises, Grundlagen der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, **5**, 52–99 (1919).
28. R. von Mises, *Wahrscheinlichkeitsrechnung, Statistik und Wahrheit*. Springer, Vienna, 1928. Second edition 1936, third 1951. A posthumous fourth edition, edited by his widow Hilda Geiringer, appeared in 1972. English editions, under the title *Probability, Statistics and Truth*, appeared in 1939 and 1957.
29. R. von Mises, *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*. Deuticke, Leipzig and Vienna (1931).
30. V. Vovk, Competitive on-line learning with a convex loss function. Working Paper No. 14, <http://www.probabilityandfinance.com> (2005).
31. V. Vovk, A. Gammerman and G. Shafer, *Algorithmic Learning in a Random World*. Springer, New York (2005).
32. V. Vovk, I. Nouretdinov, A. Takemura and G. Shafer, Defensive forecasting for linear protocols. In *AISTATS 2005: Proceedings of the 10th International Workshop on Artificial Intelligence and Statistics, Barbados, January 6–8*,

- 2005, pages 365–372, (2005). <http://www.gatsby.ucl.ac.uk/aistats/proceedings.htm>. Extended version: Working Paper No. 10, <http://www.probability-andfinance.com>.
33. V. Vovk and G. Shafer, Good randomized sequential probability forecasting is always possible. *Journal of the Royal Statistical Society, Series B*, **67**, 747–764 (2005). Extended version: Working Paper No. 7 at <http://www.probability-andfinance.com>.
34. V. Vovk, Akimichi Takemura, and Glenn Shafer. Defensive forecasting. In *Algorithmic Learning Theory: Proceedings of the 16th International Conference, ALT 2005, Singapore, October 8–11, 2005*, page 459. Springer-Verlag, 2005. Lecture Notes in Computer Science, Vol. 3734. Extended version: Working Paper No. 8, <http://www.probabilityandfinance.com>.