Probabilisitc Technology

Technology Description

Whitepaper

PredictionProbe
2192 Dupont Drive, Suite 205, Irvine, CA 92612
www.predictionprobe.com
Tel: (949) 476-4490 Fax: (949) 476-4494 info@predictionprobe.com

PredictionProbe
Probabilistic Technology Description Whitepaper

Document Purpose

The purpose of this document is to describe the core of PredictionProbe's Probabilistic Technology in layman's terms to create deeper understanding of its value.

Overview

How much time and money do you think companies spend agonizing over decisions because they cannot afford to pay for mistakes – or worse yet, they fail to make a timely decision or even to make a decision altogether? All too frequently the justification for a decision ends up being a gut feeling, a hunch, a temporary fix, the status quo, or the addition of a 30% safety factor to past estimates. In an increasingly competitive marketplace, today's companies should be very uncomfortable with such an uncertain approach to decision-making and the often unexpected and damaging results.

The ability to predict an outcome accurately gives any enterprise a tremendous advantage. Predictive Technology can revolutionize the way businesses operate, and PredictionProbe leads the way with its all-encompassing and innovative approach based on Probabilistic Technology.

The cornerstone of our Predictive Technology is **UNIPASS**™, the probabilistic software system that is being continuously enhanced and developed by our team of Probabilistic Technology experts. This software is a PC-based Windows 32 compliant FORTRAN, C++ and C software engine that can be utilized independently, as a stand-alone software engine, and/or integrated with deterministic software tools to perform complex probabilistic analysis. In the analysis, **UNIPASS**™ provides the basis for modeling uncertainties, developing probabilistic process models, computing probabilities, estimating risk, identifying most likely outcomes, providing sensitivity measures, and identifying key drivers, while the deterministic software tools may be integrated to provide the computational framework for constructing complex deterministic models. **UNIPASS**™ has available computing power that permits companies to take advantage of the full potential offered by Probabilistic Technology.

The probabilistic analysis begins by constructing process models that describe the event and the probability distribution models that explain the uncertainties associated with the process model-variables and process uncertainties. The UNIX-based process models and their associated deterministic software may be integrated with **UNIPASS™** using our proprietary **UniFace** module. We have completed development of **ModelProbe**™ **Beta Version 1.0** and **DistributionProbe**™ Version 1.2 software. These tools may be used to construct empirical process models and probability distribution models.

Using **UNIPASS**™ software and our other applications and tools, companies can quickly and accurately:

- Quantify reliability, failure probability, risk, liability, and safety measures.
- Develop optimal repair and maintenance schedules.
- Reduce maintenance and repair costs.
- Minimize inspection costs.
- Identify root-cause.
- Optimize product design.
- Reduce product development and production costs.
- Minimize warranty costs.
- Minimize product development cycle time.
- Optimize product pricing.
- Perform detailed structural analysis.
- Streamline manufacturing processes.
- Minimize weight.
- Manage data, information, and knowledge.
- Identify and mange uncertainties.
- Predict customer behavior.
- Maximize product life.
- Score and rate credit, loans or portfolios (e.g. mortgage or investment.).
- Establish valuations for commercial and residential property.
- Establish valuations for insurance claims and estimate losses.
- Develop wireless communications signal optimization and routing.
- Optimize telecommunications and network load distribution, rate maximization, speed and resources.
- Analyze and perform biotech research with improved speed and accuracy.
- Generate and optimize pricing and program models in multiple industries.
- Predict financial results, including, revenues, costs and staffing requirements, given uncertainties such as demand for products and services that normally affect these estimates.
- Allocate resources for national wilderness fire management.
- Create and enhance computer software in order to achieve the goals never realized by artificial intelligence such as voice, handwriting and visual recognition, search engines, and “smart” operating systems.
- Direct, schedule and predict bus, train, airline and other traffic, including least cost routing analysis for the transportation industry.
- Diagnose medical conditions and analyze drug enhancements.
- Establish valuations for stocks, options or futures while projecting their movements and fluctuations.
- Optimize e-commerce including customer relationship management (CRM) models, distribution channels, strategic and enterprise marketing.
- Optimize data manipulation and knowledge management resources.

These are just some of the applications and horizontal potential of the Probabilistic Technology that PredictionProbe will be able to seamlessly integrate into an organization's financial, Information Technology and other information/ management systems. PredictionProbe's software products focus on maximizing profits and eliminating waste and reducing costs from processes and business models.

Brief Technical Discussion

Traditionally, predictive analyses have used either a *deterministic* or a *statistical* approach. The deterministic approach attempts to define a process or event and then model logical outcomes using physics and/or rules, assuming all pertinent process variables to be unvarying. In contrast, the statistical approach disregards the process or event physics and/or rules and relies on pure statistical data of the process or event outcome that can often be flawed or difficult to obtain. Both of these approaches work in isolation, failing to consider any other factors that may significantly affect the result.

Probabilistic Technology incorporates the best of these two methods by modeling logical outcomes of a process or event using physics and/or rules, assuming all pertinent process variables to be random variables, i.e., as quantities which are not deterministic and which actually exhibit statistical variations. The mix is further enhanced by taking into account the uncertainties associated with the model imperfection, lack of statistical data, measurement errors, human errors, and other related uncertainties. This approach identifies the most-likely outcomes, computes the probability of the future events, and identifies the key drivers of a likely outcome; thus, providing the power to optimize the process, to zero in on the factors that actually drive the business, and to make informed decisions that will maximize business growth.

Predictive Technology Comparison Chart

PREDICTIVE METHODS	DETERMINISTIC	STATISTICAL	PROBABILISTIC
Uses physics/rule based Process Models	✓	X	✓
Uses probability law to predict outcomes	X	✓	✓
Considers various uncertainties, including inherent uncertainties, model imperfection, lack of data, human error, and more to predict outcomes	X	X	✓
Does not use probability law to construct Process Models	✓	X	✓
Does not require process output data to construct Process Models	✓	X	✓
Uses probability law to construct Input Models	X	X	✓
Compensates for unknowns using:	SF	PL	PL
Uses past performance data to improve accuracy	✓	✓	✓
Quantitatively & more accurately assesses risk, reliability, safety, performance & prediction accuracy with or without process output data	U	U	✓

✓ TRUE X FALSE U UNKNOWN SF SAFETY FACTOR PL PROBABILITY LAW

Figure 1. Comparison of Various Predictive Approaches.

Probabilistic Technology was originally developed to incorporate uncertainties into engineering processes, including, but not limited to, optimal structural analysis and design, estimating reliability of manufactured parts with limited or no performance-history, developing optimal inspection, maintenance, and repair schedules. The implementation of Probabilistic Technology will reduce design and manufacturing costs by minimizing cycle time, defects, and scrap; minimize operational costs by optimizing inspection, repair, and maintenance schedules; maintain system reliability, readiness, and effectiveness; and improve production yields, on-time delivery, and overall engineering efficiency and productivity to positively impact the bottom line. Probabilistic Technology also provides value to any business that has a need or interest in predicting the future outcome of processes or events and/or explaining why and how outcomes were obtained in the past.

Simple Functional Description

Probabilistic Technology works through the combination of the Predictive Models and our **UNIPASS**™ software engine as illustrated in Figure 2. The Predictive Models consist of Deterministic Process Models, Process Variable Models, and Uncertainty Models. A Deterministic Process Model is a mathematical representation of how an event works. It may be behavioral, process, physics and/or rule based. The Process Variable Model is a mathematical representation of the statistical behavior of the variables that enter the Deterministic Process Model to predict the outcome of the process (Process Output), including factors associated with uncertainties inherent in these variables. The Uncertainty Model is a mathematical representation of the uncertainties that could potentially influence the outcome and are not considered as part of the Variable Models. For example, a Predictive Model representing the value of a single family home may be formulated as a mathematical function of several variables including but not limited to lot size, price per square foot of the land, square footage, price per square foot of the house, number of bedrooms, number of bathrooms, location, view, etc. Many of these variables such as price per square foot of the land, price per square foot of the house, or value of the view are inherently random (i.e., quantities that are not deterministic and which actually exhibit statistical variations). Furthermore, there are other uncertainties involved in predicting the value of a single-family home such as measurement error, appraiser's bias, or the mathematical function that predicts the value of such home. Therefore, a more accurate valuation may be reached by incorporating such uncertainties—i.e., using Probabilistic Technology. The **UNIPASS**™ software engine uses the Predictive Models and produces a host of vital information that allows decision makers to reach their best decisions given the available information.

An Example of Probabilistic Technology

New ideas are often best explained using simple examples. Here is a simple example of how Probabilistic Technology works. Let's say, that you wanted to figure out how long it is likely to take to get from PredictionProbe Headquarters in Irvine, CA to Los Angeles International Airport (LAX).

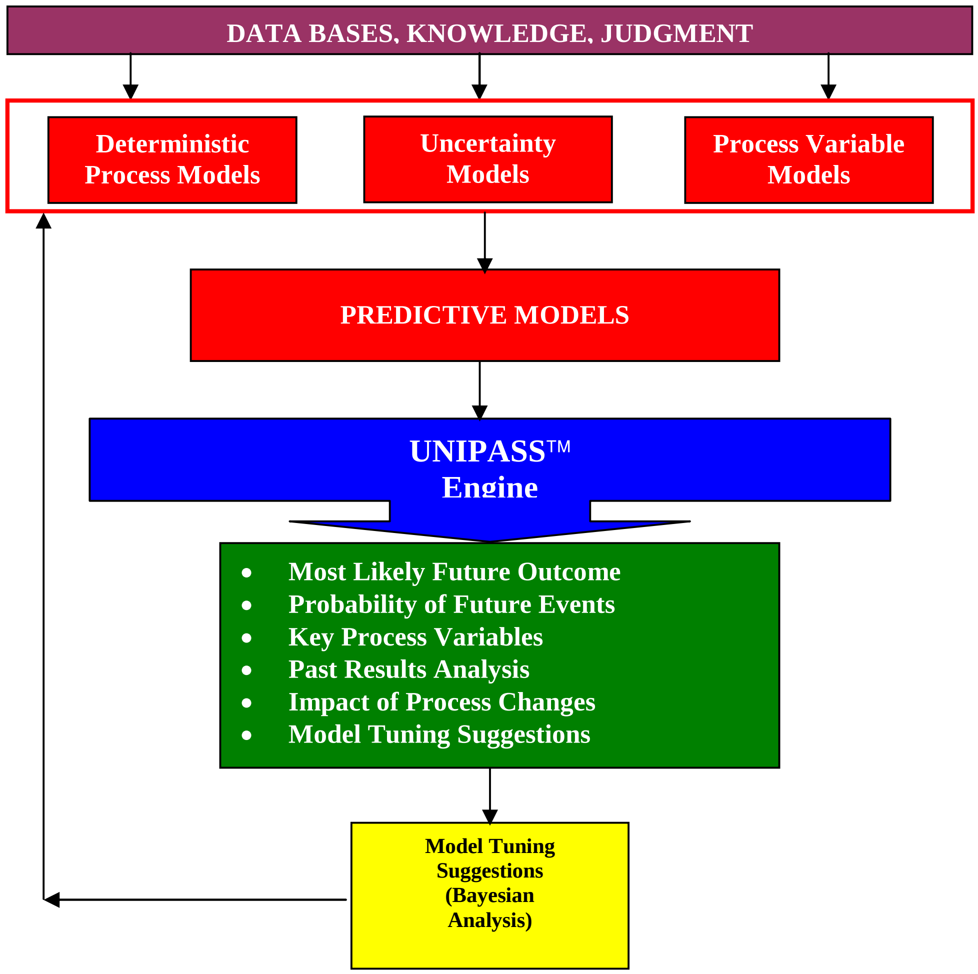

Figure 2. Simple Illustration of How Probabilistic Technology Works.

Assume that you are going to LAX from PredictionProbe Headquarters in Irvine, CA on Monday morning at 10:00 a.m. You are going to take the most direct route (the I-405 Freeway) that is a distance of 45 miles. CalTrans also has data that provides the range of speed for the I-405 Freeway from Irvine to LAX for every day of the week and is recorded hourly. For Monday morning at 10:00 a.m., the speed on this stretch of the I-405 freeway averages 45 mph and ranges from 35 mph to 55 mph. You want to predict with a high degree of accuracy how long it will take you to make this trip because you have important meetings to attend at PredictionProbe early on Monday morning and would like to stay as late as possible, but you also want to have only a small probability of missing your flight and would like to leave early enough to avoid any mishaps.

Like any other predictive problem, there are three approaches you can use: Deterministic, Statistical, or Probabilistic.

Deterministic Approach

A deterministic approach allows you to use physics to establish your answer. The calculation is simple: time = distance/speed. Since the distance is known and is fixed at 45 miles (namely, LAX and PredictionProbe Headquarters don't move), the formula simplifies to: time = 45/speed. The only unknown variable is speed. The deterministic approach uses a value for speed (in some instances, it uses the average). In this example, that would be 45 mph. As a result, we'd say that we'd expect it to take 45/45 = 1 hour to get to the airport. However, the deterministic approach also assigns "safety factors" in order to account for overlooked variability—in this case, we could say that a 25% safety factor could be applied and we'd predict we should allow 1 hour + 25% or 1.25 hours to get to the airport. Sometimes, the deterministic approach would apply a safety factor on the speed. For example, we could use the slowest speed (35 mph) and calculate time as 45/35 = 1.29 hours. Sometimes, the deterministic approach would apply a safety factor to both the speed and the final time. In this example, that would put the time at 1.29 + 25% or 1.61 hours.

Of course, there are big problems with this approach. First, it seems awfully ad-hoc. Why a safety factor of 25% and not of 20% or 30% or 50%. It provides no quantification of how likely it is that you'll make the flight on time or not. If you use the slowest speed and 25% safety factor, you may never miss a flight, but it seems like you'd be leaving much too early on most occasions. Clearly, this prediction method is not precise enough for this problem.

Statistical Approach

A statistical approach allows you to compensate for unknowns in a better fashion by collecting a statistically significant sample of data and making decisions based on it. Normally, a statistically significant sample would be between 50 to 100 data points. In this example, we would need to collect times of 50 to 100 drivers going from PredictionProbe Headquarters to LAX on Monday morning at 10:00 a.m. If we were able to collect these data accurately, we could fit some statistical model to help predict how long it would take to drive. In Figure 3 below, the histogram shows 100 times which were collected and a statistical model (bell curve or normal distribution) that will predict any probability that is of interest. For example, this statistical model predicts that it will take more than 1.16 hours only 5% of the time (since 5% of the area under the curve is to the right of 1.16 hours). That is useful information as we can now quantify the trade-off between leaving too early and missing our flight.

But a glaring problem presents itself here. Namely, the statistical approach assumes that these very special data either already exists somewhere or that we will collect them. This data does not presently exist and thus we would be required to go and collect this information. This would no doubt be quite expensive and time consuming. Sometimes people collect statistical data but there are many problems with such data. For example, we could probably collect 50 or 100 data points if we worked hard enough by asking various people who travel a lot, but we would likely have drivers traveling on days other than Monday, times other than 10:00 a.m., taking routes other than the I-405 Freeway, driving in a variety of different weather conditions, etc. As a result, without a very specific controlled study, our results would not be too useful. But, fortunately there's a better way.

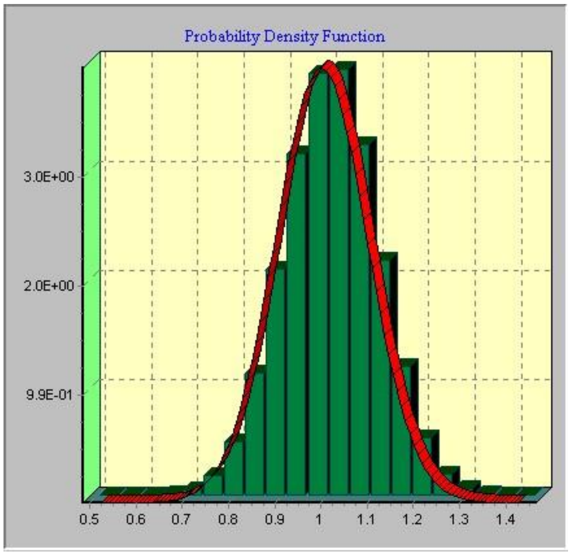

Figure 3. Distribution of Time to Airport (in hrs)

Using Statistical Approach

Probabilistic Approach

A probabilistic approach to this problem combines the best of the deterministic and statistical approaches and provides an extremely high degree of precision in the prediction. Probabilistic Technology allows you to model the process of driving to the airport using the advantages of *both* deterministic and statistical approaches. In this example, we use the physical process model (time = 45 / speed) that the deterministic approach used but account for variation in speed using a statistical approach. In particular, if we plot the data from CalTrans that shows the speed range of 35 mph to 55 mph (see Figure 4), we can develop a distributional model (skewed curve on the histogram) to best describe the variation in speed. Speed is thus our Input Variable, and the distribution in Figure 4 accounts for the Uncertainty in Speed. Note that we use data that are readily available—not data that are difficult or expensive to obtain.

Next, the probabilistic approach uses the physical model in combination with this variation in speed to predict the variability in the time it will take to get from PredictionProbe Headquarters to LAX on Monday at 10.00 a.m. The result is illustrated in Figure 5.

Note we get the advantages of both the deterministic and the statistical approaches. We use the actual physical model, which is the benefit of using the deterministic approach. Likewise, we avoid having to use ad-hoc safety factors by considering variability accurately, which is the benefit of using the statistical approach. In this example, using the probabilistic approach, we conclude that 5% of the time, it will take us longer than 1.08 hours to get to LAX. Likewise, we can see that it would take us longer than 1.16 hours only 1% of the time. This allows us to trade-off risk (probability of missing the flight) and cost (how much time we're going to allocate to get to the airport and not be able to work at PredictionProbe).

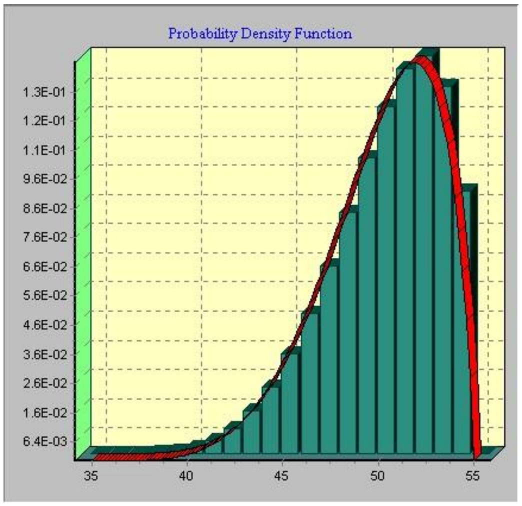

Figure 4. Distribution of Speed (in mph) at 10.00 a.m.. on Monday from Irvine, CA to LAX

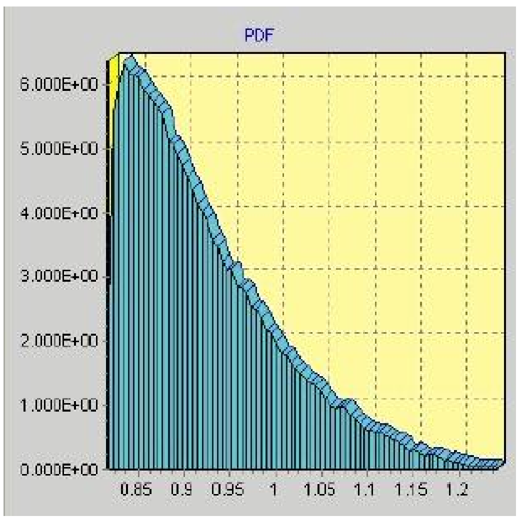

Figure 5. Distribution of Time (in hrs) to Airport Using Probabilistic Approach.

More Complex Applications of Probabilistic Technology

As you might expect, the example is very simple in terms of illustrating the concepts of Probabilistic Technology whereas many applications using the technology are quite complex. For example, the Deterministic Process Model we used was very simple and we assumed the same mathematical model for both Predictive Model and Deterministic Process Model. Some Deterministic Process Models would be a single complex equation, others might be a series of complex equations, and still others would require the execution of complex computer codes to obtain results. Also, our Deterministic Process Model had only one Process Variable, speed, and one Process Output, time. Deterministic Process Models often times have numerous Process Variables and many Process Outputs. But the bottom line is that the underlying concept of a Probabilistic analysis is the same whether the individual components are simplistic as in our example or complex as in many applications.

If our "drive to the airport" example above were more complex, the Predictive Model may have included variables that are not used in the Deterministic Process Models such as weather conditions, whether or not there was an accident, the month of the year, the impact of a simplified Deterministic Process Model, etc. These would certainly result in a more accurate model than just using the simpler time = distance/speed relationship. If we did develop a time as a function of all of these variables relationship, several benefits of the probabilistic approach could be obtained:

- The key variables causing different times could be identified and ranked from most important to least important (e.g., whether or not there is an accident is most important, weather conditions are second, etc.).
- You could identify the reason for late arrival of a driver to LAX (e.g., it's raining, there is an accident, or it's the first day of school).
- You can determine the most accurate range based on the given information (e.g., based on today's conditions, there is a 90% chance that it will take between 53 minutes and 1 hour and 4 minutes to get to LAX).
- You can evaluate different drivers' performance against the model to determine its accuracy and fine-tune it for even more accurate future predictions. You could even use it to "flag" unusual events. For example, someone always arriving too early might be consistently excessively speeding or someone always missing his flight may always be making a stop at the store first.

It is impossible to determine these findings using the deterministic or statistical approach.

Extensions To Various Industries

In the simple example above, the concepts of a Probabilistic approach were illustrated. This approach can be extended to a large number of products in large, untapped, and rapidly growing markets. These markets include the majority of the world's leading businesses, operating in multiple industries such as engineering, insurance, real estate, medical, pharmaceutical, financial, investment markets, manufacturing, biotechnology, telecommunications, wireless technology, transportation, e-commerce and many others. The following sections briefly discuss the potential extensions to the Engineering and Insurance industries.

Extensions To Engineering Problems

The Probabilistic approach demonstrated above could be extended to those problems that concern Engineering. The methodology that is used is no different. Only the Process Variables, Process Output, and Deterministic Process Models are different. Instead of time to get to the airport, Engineers may be concerned with predicting the stresses on their parts or the product life or reliability of a design. Instead of distance and speed, engineers might be concerned with Process Variables such as temperatures, geometries, tolerances, material properties, flows, etc. Likewise, engineering Deterministic Process Models are mathematical expressions that may be constructed using the well-defined equations with numerous variables, obtained from commercial or proprietary engineering handbooks. These mathematical expressions are often much more complex than the time = distance/speed example we used earlier. If well-defined equations are not available, these models may be constructed using a complex software tool. For example, ANSYS finite element software is one such software. This software will construct complex structural models that calculate the stress based on numerous Process Variables such as geometries, material properties, loads, temperatures, etc.

By integrating process physics and various uncertainties Probabilistic Technology allows for optimal product design, accurate product life prediction, quantitative risk assessment, product reliability and safety improvements, minimal warranty costs, optimal inspection and repair planning, life-cycle-costs optimization, optimal manufacturing process control, advanced six-sigma implementation, and more. This can benefit a variety of industries that design, test, construct, maintain, and operate various products. Improved safety assurance is a goal of the nuclear power industry. Controlled maintenance costs and longer life are required for bridges, offshore oil drilling platforms, and ocean-going ships. The automotive and ground transportation industries need to reduce costs associated with warranty, high fuel consumption, and high maintenance requirements. In the aerospace industry, weight reduction and reduced maintenance lead to reduced payload transportation costs. A more realistic method of designing buildings to make them earthquake resistant is needed by the construction industry. Reducing costs through anticipating and eliminating failures is a goal of many industries. Incorporating Probabilistic Technology into product design criteria is a good way to benefit all of these industries.