RISK PREDICTION FOR MODERN TECHNOLOGICAL SYSTEMS

Duffey Romney B.
Atomic Energy of Canada Limited,
Chalk River, ON, Canada

Saull John W.
International Federation of Airworthiness,
East Grinstead, UK

Keywords
technological systems, risk, outcome, failure, error, events, probability

Abstract
We have already examined the worldwide trends for outcomes (measured as accidents, errors and events) using data available for large complex technological systems with human involvement. That analysis was a dissection of the basic available, published data on real and measured risks, for trends and inter-comparisons of outcome rates. We found and showed how all the data agreed with the learning theory when the accumulated experience is accounted for. Here, learning includes both positive and negative feedback, directly or indirectly, as a result of prior outcomes or experience gained, in both the organizational and individual contexts. Our purpose here and now is to try to introduce some predictability and insight into the risk or occurrence of these apparently random events. In seeking such a general risk prediction we adopt a fundamental theoretical approach that is and must be testable against the world's existing data. Comparisons with outcome error data from the world's commercial airlines, the two shuttle failures, and from nuclear plant operator transient control behaviour, show a reasonable level of accord. The results demonstrate that the risk is dynamic, and that it may be predicted using the MERE learning hypothesis and the minimum failure rate, and can be utilized for predictive risk analysis purposes.

1. The risk prediction purpose

Modern technological systems fail, sometimes with catastrophic consequences, sometimes just everyday injuries and deaths. The risk is given by the probability of failure, error or more generally any outcome. Recently the crash of the NASA Space Shuttle *Columbia*, the great blackout of the North East USA and Canada, the explosion at the Texas City refinery all occurred. Other smaller but also key accidents have also occurred: the midair collision over Europe of two aircraft carrying the latest collision avoidance system; the glider landing of a jet aircraft out of fuel in the Azores; a concrete highway overpass collapsing in Laval, Quebec; the huge oil tank fire in England; more ships sinking, more trains derailing, even more cars colliding, and evermore medical errors. We have already examined the worldwide trends for outcomes (measured as accidents, errors and events) using data available for large complex technological systems with human involvement. That analysis was a dissection of the basic available, published data on real and measured risks, for trends and inter-comparisons of outcome rates. We found and showed how all the data agreed with the learning theory when the accumulated experience is accounted for. Here, learning includes both positive and negative feedback, directly or indirectly, as a result of prior outcomes or experience gained, in both the organizational and individual contexts as in [5]. Our purpose here and now is to try to introduce some predictability and insight into the risk or occurrence of these apparently random events. In seeking such a general risk prediction we adopt a fundamental theoretical approach that is and must be testable against the world's existing data.

2. What we must predict

We have shown how outcomes develop in phases from a string or confluence of factors too complex to predict but always avoidable. The bright feature is that we now know that a universal learning curve (ULC) exists and we can utilize that to predict outcome rates and track our progress as we improve. We can therefore start to manage the risk, but only if we include the human element.

We need to make it entirely clear what we do not propose. We will not use the existing idea of analysing human reliability and errors on a task-by-task, item-by-item, situation-by-situation basis. In that approach, which is commonly adopted as part of probabilistic safety analysis using event sequence "trees", the probability of a correct or incorrect action is assigned at each significant step or branch point in the hypothesized evolution of an accident sequence. The probability of any action is represented and weighted or adjusted by situational multipliers, representing stress, environment and time pressures. We suggest, at least for the present, that it is practically *impossible* to try to describe all the nuances, permutations and possibilities behind human decision-making. Instead, we treat the homo-technological system (HTS) as an integral system. We base our analysis on the Learning Hypothesis, invoking the inseparability of the human and the technological system. Using the data, we invoke and use experience as the correct measure of integrated learning and decision-making opportunity; and we demonstrate that the HTS reliability and outcome probabilities are dynamic, simply because of learning.

The basic and sole assumption that we make every time and everywhere is the "learning hypothesis" as a physical model for human behaviour when coupled to any system. Simply and directly, we postulate that humans learn from their mistakes (outcomes) as experience is gained. So, the rate of reduction of outcomes (observed in the technology or activity as accidents, errors and events) is proportional to the number of outcomes that are occurring.

That learning occurs is implicitly obvious, and the reduction in risk must affect the outcome rate directly. To set the scene, let us make it clear that the probability of error is quite universal, and can affect anyone and everyone in a homo-technological system (HTS). There are clear examples of highly skilled well-trained operators, fully equipped with warning and automated systems. So all the people involved (from maintenance, ground control, management, airline operator and the pilots) are working in an almost completely safe industry (ACSI).

Two aircraft examples are very basic to safety: loss of fuel while in flight, and mid-air collision. Given all the systems put in place to avoid these very obvious and fundamental risks, the outcomes still occurred. But despite all the effort, procedures and warnings, there is loss of control through loss of understanding, communication and information in the most modern of aircraft which were maintained to the highest standards. We need to estimate their chance of occurrence of the outcomes, and define the risk by finding the probability of the outcomes due to the human errors embedded in the HTS.

Let us start with the learning hypothesis applied and applicable to any integrated (total) HTS. Thus, the human error or technological system failure or outcome rate, λ, is equivalent to a dynamic hazard function $h(\varepsilon)$ which varies with experience, ε, as given by the Minimum Error Rate Equation (MERE):

$$d\lambda/d\varepsilon = - k (\lambda - \lambda_m) \quad (1)$$

where k is the learning rate constant, and λ_m the minimum obtainable rate. The failure or outcome rate as a function of experience, $\lambda(\varepsilon)$, is then obtained by straightforward integration as,

$$\lambda(\varepsilon) = \lambda_m + (\lambda_0 - \lambda_m) e^{-k\varepsilon} \quad (2)$$

where the outcome or failure rate $\lambda \equiv h(\varepsilon)$, the hazard function; λ_m is the minimum obtainable rate at large experience; and λ_0 is the initial rate at some initial experience, ε0.

Here, it will be remembered that the failure or outcome rate is the summation of all the ith rates in the technological system, so that effectively:

$$\lambda(\varepsilon) - \lambda_m = \Sigma_i (\lambda_i - \lambda_m) \quad (3)$$

Since the MERE result describes and agrees with a wide range of actual data, we hypothesize that this is indeed the correct form for the human error or outcome rate in a HTS with learning. This form has been used to derive the ULC, validated by obtaining failure rates from the world accident, injury and event data.

In terms of the number of failures, errors or observed outcomes, N_j, then we have when sampling the j^{th} observation interval the hazard function or failure rate:

$$\lambda(\varepsilon) = \{(1/(N - N_j))\ (dN_j/d\varepsilon)\} \tag{4}$$

where N is the total number of outcomes and $A \equiv dN_j/d\varepsilon$, the instantaneous outcome rate, IR, and the number of outcomes we have observed over all prior intervals is just the summation, $n \equiv \Sigma_j N_j$.

3. The probability linked to the rate of errors

Given the outcome rate, now we need to determine the outcome (error) probability, or the chance of failure.

a. the hazard function is equivalent to the *failure or outcome rate* at any experience, $\lambda(\varepsilon)$, being the relative rate of change in the reliability with experience, $1/R(\varepsilon)\ (dR(\varepsilon)/d\varepsilon)$;
b. the *CDF or outcome fraction*, $F(\varepsilon)$, is just the observed frequency of prior outcomes, the ratio n/N, where we have recorded n, out of a total possible of N outcomes;
c. the *frequency of prior outcomes* is identical to the observed *cumulative prior probability*, $p(\varepsilon)$, and hence is the CDF, so $F(\varepsilon) = p(\varepsilon) = (n/N) = 1 - R(\varepsilon)$;
d. here $R(\varepsilon)$ is the *reliability,* 1-n/N, a probability measure of how many outcomes or failures did *not* occur out of the total;
e. the *future (or Posterior) probability*, p(P) is proportional to the Prior probability, $p(\varepsilon)$ times the Likelihood, p(L), of future outcomes;
f. the chance of an outcome in any small observation interval, is the PDF $f(\varepsilon)$, which is just the rate of change of the failure or outcome fraction with experience, $dp(\varepsilon)/d\varepsilon$;
g. the *Likelihood,* p(L) is the ratio, $f(\varepsilon)/F(\varepsilon)$, being the probability that an outcome will occur in some interval of experience, the PDF, to the total probability of occurrence, the CDF; and
h. we can write the PDF as related to the failure rate integrated between limits from the beginning with no experience up to any experience, ε,

$$f(\varepsilon) = dF/d\varepsilon = \lambda(\varepsilon) \exp - \int_0^{\varepsilon} \lambda(\varepsilon)\, d\varepsilon. \tag{5}$$

So, the probability of the outcome or error occurring in or taking less than ε, is just the CDF, $p(\varepsilon) = n/N$, conventionally written as $F(\varepsilon)$. Relating this to the failure rate, via (a) through (d) above, gives:

$$p(\varepsilon) \equiv F(\varepsilon) = 1 - e^{-\int \lambda d\varepsilon} \tag{6}$$

where, of course from the MERE,

$$\lambda(\varepsilon) = \lambda_m + (\lambda_0 - \lambda_m) \exp - k(\varepsilon - \varepsilon_0) \tag{7}$$

and $\lambda(\varepsilon_0) = \lambda_0$ at the initial experience, ε_0, accumulated up to or at the initial outcome(s). The corresponding PDF $f(\varepsilon)$, is the probability that the error or outcome occurs in the interval dε, derived from the change in the CDF failure fraction with experience, or from (f), (h) and (g) above:

$$f(\varepsilon) = dF(\varepsilon)/d\varepsilon = dp(\varepsilon)/d\varepsilon = \lambda e^{-\int \lambda d\varepsilon} = \lambda(\varepsilon) \text{ x } (1 - p(\varepsilon))$$

$$= \{(\lambda_m + (\lambda_0 - \lambda_m) \exp(-k(\varepsilon - \varepsilon_0))\} \text{ x } \{\exp((\lambda(\varepsilon) - \lambda_0)/k - \lambda_m(\varepsilon_0 - \varepsilon))\} \tag{8}$$

The limits are clear: as experience becomes large, $\varepsilon \rightarrow \infty$, or the minimum rate is small, $\lambda_m << \lambda_0$, or the value of k varies, etc. We can also show that the uniform probability assumption for *observing* outcomes is consistent with the systematic variation of the outcome probability with experience due to *learning*.

We can also determine the maximum and minimum risk likelihood's, which are useful to know, by differentiating the probability expression. The result shows how the risk rate systematically varies with experience and that the most likely trend is indeed given by the learning curve. In other words, we learn as we gain experience, and then reach a region of essentially no decrease, in rate or in probability, and hence in likelihood. It is easy to obtain the first decrease in rates or probabilities but harder to proceed any lower. This is exactly what is observed in transport, manufacturing, medical, industrial and other accident, death and injury data [2].

4. The initial failure rate and its variation with experience

Having established the learning trend, we need to determine the actual parameters and values using data and insight. Now, in reality, the initial rate, λ_0, is *not* a constant as assumed so far since the outcomes are stochastic in experience "state space". Hence, $\lambda_0 = \lambda(\varepsilon_0)$, and it is not known when exactly in our experience we may have an error initially observed (and we might be lucky or not), and the initial value we ascribe to the initial rate observe is an arbitrary value.

To establish the initial rate, key data are available from commercial aircraft outcomes (fatal crashes) throughout the world. The major contributor is human error not equipment failure, although these latter can also be ascribed to the root cause of human failings. Fatal crashes and accidents for the thirty years between 1970 and 2000 are known [1], for 114 major airlines with ~725 million hours (Mh) of total flying experience. For each airline with its own experience, ε, the fatal accident rate per flying hour, $\lambda(\varepsilon)$, can be plotted as an open circular symbol in Figure 1 versus the accumulated experience in flying hours (adopting the FAA value of ~31/3 hours as an average flight time).

These are:

a) the crash of the supersonic *Concorde* with a rate, λ_0, of one in about 90,000 flights shown as a lozenge symbol; and
b) the explosion and disintegration of the space shuttles, *Challenger* and *Columbia,* with a rate, λ_0, of two out of 113 total missions, plotted using

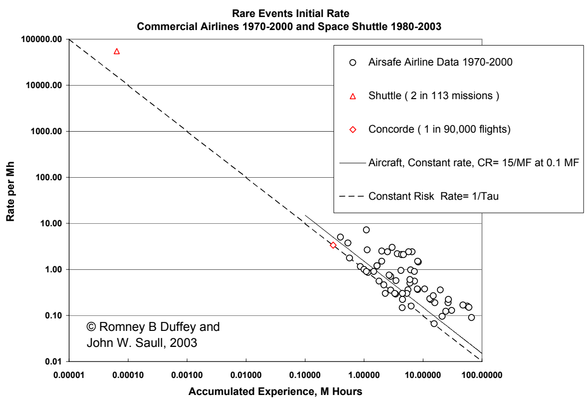

Figure 1. The initial rate based on world airline and US space shuttle accident data

the triangular symbol. The typical "flight time" for the shuttle was taken as the 30-40 minutes for re-entry as reported by NASA [4] timelines, although this plot is quite insensitive to the actual value taken.

For all these data and experience, there is a remarkable constancy of risk, as shown by the straight line of slope –1, which is given by the equation:

$$\lambda\varepsilon = \text{constant, n,} \tag{9}$$

where the observed rate is strictly a function of whatever experience it happened to occur at, any value being possible. Thus, in the limit for *rare events,* the initial rate should be the purely Bayesian estimate from the prior experience with n ~1 and $\lambda_0 \approx (1/\varepsilon)$. This rate varying as $(1/\varepsilon)$ also corresponds *exactly* to the risk rate that is attainable on the basis of the minimum likelihood determined from the outcome probability. What the data are telling us is that the limiting initial rate is exactly what it is for the experience at which the first outcome occurs, no more and no less.

From the analysis of many millions of data points that include human error in the outcomes, we have been able to derive the key quantities that dominate current technological systems. These now include commercial air, road, ship and rail transport accidents; near-misses and events; chemical, nuclear and industrial injuries; mining injuries and manufacturing defects; general aviation events; medical misadministration and misdiagnoses; pressure vessel and piping component failures; and office paperwork and quality management systems [2].

From all these data, and many more, we have estimated the minimum failure rate or error interval, the typical initial error interval, and the learning rate constant for the ULC as follows:

a) minimum attainable rate, λ_m, at large experience, ε, of about one per 100,000 to 200,000 hours ($\lambda_m \sim 5.10^{-6}$ per hour of experience);
b) initial rate, λ_0, of $1/\varepsilon$, at small experience (being about one per 20,000 to 30,000 hours or $\lambda_0 \sim 5.10^{-5}$ per hour of experience);
c) learning rate constant, $k \sim 3$, from the ULC fit of a mass of available data worldwide for accidents, injuries, events, near-misses and misadministration.

Therefore, the following numerical dynamic form for the MERE human error or outcome rate is our "best" available estimate [2]:

$$\lambda(\varepsilon) = \lambda_m + (\lambda_0 - \lambda_m)\, e^{-k\varepsilon}, \tag{10}$$

which becomes, for $\lambda_0 = (n/\varepsilon)$, with n = 1 for the initial outcome,

$$\lambda = 5.10^{-6} + (1/\varepsilon - 5.10^{-6})\, e^{-3\varepsilon} \tag{11}$$

The rate, λ, can be evaluated numerically, as well as the probability, $p(\varepsilon)$, and the differential PDF, $f(\varepsilon)$. The result of these calculations is shown in *Figure 2*, where $\varepsilon \equiv \tau$ units in order to represent the accumulated experience scale.

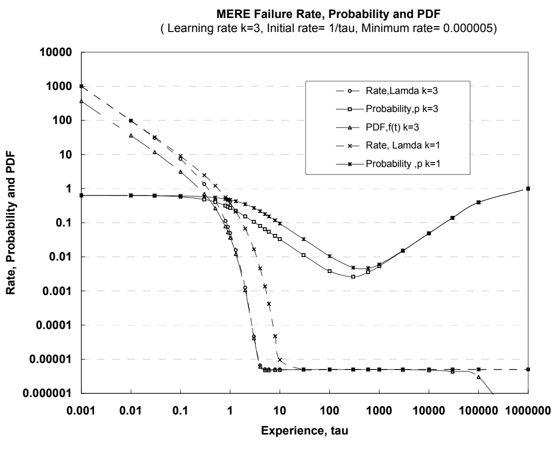

Figure 2. The best MERE values

It is evident that for k>0 the probability is a classic "bathtub" shape, being just under near unity at the start (*Figure 2*), and then falling with the lowering of error rates with increasing experience. After falling to a low of about one in a hundred "chance" due to learning, it rises when the experience is $\varepsilon > 1000$ tau units, and becomes a near certainty again by a million tau units of experience as failures re-accumulate, since $\lambda_m \sim 5.10^{-6}$ per experience tau unit. The importance of learning is evident, since for k<0 forgetting causes a rapid increase to unity probability with no minimum. The solution for the maximum likelihood for the outcome rate is exponential, falling with increasing experience as given by: (Rate for Maximum Likelihood)

$$\lambda_{max} = \lambda_0 \exp - \{k(\varepsilon - \varepsilon_0)/(1 + k\varepsilon_0)\} \tag{12}$$

However, the expression that gives the minimum likelihood indicates that the *minimum risk* rate is bounded by: (Rate for Minimum Likelihood)

$$\lambda_{min} << \{ \lambda_m \}/ \{1 + k(\varepsilon - \varepsilon_0) \} \quad (13)$$

The result follows common sense. Our maximum risk is dominated by our inexperience at first, and then by lack of learning, and decreasing our risk rate largely depends on attaining experience. Our most likely risk rate is extremely sensitive to our learning rate, or k value, for a given experience.

So, as might be logically expected, the *maximum likelihood for outcomes occurs at or near the initial event rate when we are least experienced.* This is also a common sense check on our results: *we are most at risk at the very beginning.* Therefore, as could have been expected, the most likely and the least risks are reduced only by attaining increasing experience and with increased learning rates.

This approach to reduce and manage risk should come as no surprise to those in the education community, and in executive and line management positions. *A learning environment has the least risk.*

5. Future event estimates: the past predicts the future

The probability of human error, and its associated failure or error rate, we expect to be unchanged unless dramatic technology shifts occur. We can also estimate the likelihood of another event, and whether the MERE human error rate frequency gives sensible and consistent *predictions*. Using Bayesian reasoning, the posterior or future probability, p(P), of an error when we are at experience, ε, is,

$$\text{Posterior, } p(P) \propto \{\text{Prior, } p(\varepsilon)\} x \{\text{Likelihood, } p(L)\} \quad (14)$$

where p(ε) is the prior probability, and by definition both |P,L| > ε, our present accumulated experience**.** The likelihood, p(L), is also a statistical estimate, and we must make an assumption, based on our prior knowledge, and often is taken as a uniform distribution. We can show that the likelihood is formally related to the number of outcomes for a given variation of the mean.

Either:

a) the future likelihood is of the same form as experienced up to now; and/or
b) the future is an unknown statistical sample for the next increment of experience based on the differential probability, the PDF f(ε).

In the first case (a), we have that the future likelihood probability p(L) is the fraction or ratio of events remaining to occur out of the total possible number that is left. For the second case (b), the future is an unknown statistical sample for the next increment of experience based on the PDF, f(ε). This is called a "conditional probability", where the probability of the next outcome depends on the prior ones occurring, which was Bayes original premise.

The so-called generalized *conditional* probability or Likelihood, p(L), can be defined utilizing the CDF and PDF expressions. Described by [6] as the "generalized Bayes formula", the expression given is based on the prior outcome having already occurred with the prior probability p(ε). This prior probability then gives the probability or Likelihood of the next outcome, p(L), in our present experience-based notation, as:

$$p(L) = \equiv \text{PDF/CDF}\} \equiv = \lambda\{(1-p(\varepsilon))/p(\varepsilon)\} \quad (15)$$

We can evaluate this Bayesian likelihood and posterior expressions using our "best" MERE values of a learning rate constant of k=3 and a minimum failure rate of $\lambda_m = 5.10^{-6}$, obtaining the results shown in *Figure 3.*

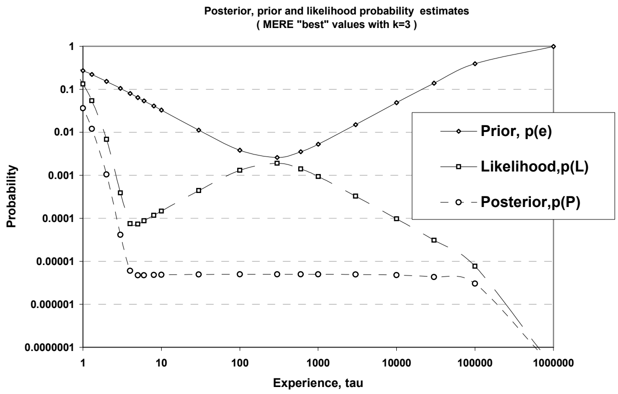

Figure 3. The estimate of the likelihood and posterior probabilities when learning

It is clear from *Figure 3* that the "human bathtub" prior probability, p(ε), causes the likelihood to fluctuate up and down with increasing experience. The likelihood tracks the learning curve, then transitions via a bump or secondary peak to the lowest values as we approach certainty ($p \rightarrow 1$) at large experience. However, the posterior probability, p(P), just mirrors and follows the MERE failure rate, as we predicted, decreasing to a minimum value of $\sim 5.10^{-6}$, our ubiquitous minimum outcome rate, before finally falling away.

Hence, since the future probability estimate, the posterior p(P), is once again derivable from its (unchanged) prior value, $f(\varepsilon) = dp(\varepsilon)/d\varepsilon \sim \lambda(\varepsilon)$, derived from learning from experience, and thus *the past predicts the future*.

For the special case of "perfect learning" when we learn from all the non-outcomes as well as the outcomes, the Poisson-type triple exponential form applies for low probabilities and small numbers of outcomes ($n \ll m$). Of course, the limit of "perfect learning" is when we have an outcome, so here $p(\tau) = 1/\tau$, and is the rare event case for n = 1. The Perfect Learning limit fails as soon as we have an event, as it should. But there is also a useful simple physical interpretation, which is that:

a) we learn from non-outcomes the same way we learn from outcomes, as we have assumed;
b) the perfect learning ends as soon as we have just a single (rare) outcome; and
c) the influence of the finite minimum rate is then lost.

6. Comparison to data: the probability of failure and human error

There are three data sets for catastrophic events with defined large human error contributions that are worth re-examining further:

a) the crash rate for global commercial airlines, noting most occur during manoeuvring and approach for take-off and landing but as we have seen can also occur in flight;
b) the loss of the space shuttles, *Challenger* and *Columbia*, also on take-off and during the approach for landing; and
c) the probability of non-detection by plant operators of so-called latent (hidden) faults in the control of nuclear system transients.

Apparently disparate, these three all share the *common element of human involvement* in the management, design, safety "culture", control and operation of a large technological system; all are subject to intense media and public interest; and the costs of failure are extremely expensive and unacceptable in many ways.

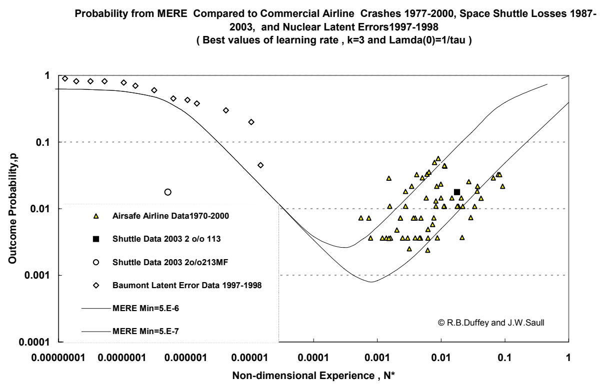

Figure 4. An outcome probability data comparison

The comparison of the data to theory is shown in *Figure 4* where the lines are the MERE calculated probability, p(ε) using the "best" values. The three lines use three bounding values for the minimum error rate to illustrate the sensitivity. Despite the scatter, a minimum rate of order $\sim 5.10^{-6}$ is indeed an upper bound value, as we estimated before.

7. Implications for generalized risk prediction

The implications of using this new approach for estimating risk are profound.

This new probability estimate is based on the failure rate describing the ULC, which is derived from the Learning Hypothesis; and utilizes the validation from the outcome data of the world's homo-technological systems. Thus, we have seamlessly linked all the way from individual human actions to the observed outcomes in entire systems. We have unified the approach to managing risk and error reduction using the Learning Hypothesis with the same values everywhere for the learning rate constant, k, and the minimum error rate, λ_m.

For the first time, we are also able to make predictions of the probability of errors and outcomes for any assumed experience interval in any homo-technological system.

Typically the probabilities for error are $\sim 10^{-2}$, or one in a few hundred, for any act of volition beyond the first 10% or so of the risk interval; whereas in that first increment or initial phase, the risk is much higher. Interestingly, this reduction in probability in the initial interval echoes, parallels and is consistent with the maze study results of Fang [3] showed a rapid (factor of five or so) reduction in the first 10% or so of the moves needed for success by the "treasure hunt" players. Clearly the same fundamental learning factors and success motivation is at work, and are reflected in the rapid decrease in errors down the learning curve.

Conversely, the MERE probability (the human bathtub) properly represents the data trends, such as they are, and hence can be used in PRA HEP estimation provided the correct measure is taken for experience.

In an addition the MERE results implies a finite lower bound probability of $p(\varepsilon) > 10^{-3}$, based on the best calculations and all the available data.

8. Conclusions: the probable risk

Analysis of failure rates due to human error and the rate of learning allow a new determination of the risk due to dynamic human error in technological systems, consistent with and derived from the available world data. The basis for the analysis is the "learning hypothesis" that humans learn from experience, and consequently the accumulated experience defines the failure rate. The exponential failure rate solution of the Minimum Error Rate Equation defines the probability of human error as a function of experience.

Comparisons with outcome error data from the world's commercial airlines, the two shuttle failures, and from nuclear plant operator transient control behaviour, show a reasonable level of accord. The results demonstrate that the risk is dynamic, and that it may be predicted using the learning hypothesis and the minimum failure rate, and can be utilized for predictive risk analysis purposes.

References

[1] Airsafe (2000). Fatal Events and Fatal Event Rates from 1970-2000, September, http://www.airsafe.com.

[2] Duffey, R.B. & Saull, J.W. (2002). *Know the Risk.* First Edition, Butterworth and Heinemann, Boston, USA.

[3] Fang, Christina. (2003). Stern School of Business, New York. Learning in the absence of feedback, unpublished MS.

[4] NASA. Implementation Plan for Return to Flight, National Aeronautics and Space Administration, http://www1.nasa.gov).

[5] Ohlsson, S. (1996). Learning from Performance Errors. *Psychological Review*, Vol. 103, No. 2, 241-262.

[6] Sveshnikov, A.A. (1968). Problems in Probability Theory, Mathematical Statistics and the Theory of Random Functions. Dover, New York.