JOURNAL OF SPACECRAFT AND ROCKETS
Vol. 41, No. 1, January–February 2004

# Bayesian Analysis of Launch Vehicle Success Rates

Seth D. Guikema* and M. Elisabeth Paté-Cornell†
*Stanford University, Stanford, California 94305-4026*

**In the choosing of a launch vehicle for a given mission or in the determination of insurance coverage and premiums for a given launch, accurate estimates of the probability of success of the different launch vehicles provide important information. There are three general approaches to estimating the probability of launch success. The first is to use a probabilistic risk analysis, decomposing the system into its subsystems and components and estimating the probability of each of the failure modes. The second is to rely on expert judgment about the vehicle's success rate as a whole, without a functional decomposition of the system. The third is to use statistical data about the past performance of the system to estimate the vehicle's success rate. The focus is put on this last approach, using Bayesian probability theory to make better use of vehicle-level performance data. The procedure is demonstrated by an analysis of the success rates of most of the major families of launch vehicles currently in use in the world. A family of launch vehicles includes all variants of a particular type of vehicle from a specific manufacturer, for example, the Delta 2. For vehicles with a small number of launch attempts, the Bayesian approach provides the advantage over classic statistical approaches of yielding estimates of both the mean future frequency of success and the uncertainty about that mean.**

## Nomenclature

$A$ = future frequency of launch success for a given launch vehicle, random
$a$ = realized success rate for a given vehicle
$D$ = launch history (data) of a given launch vehicle, consisting of $s$ successes in $t$ trials
$s$ = number of successes a launch vehicle has had in its $t$ trials
$t$ = number of past launch attempts with a given launch vehicle

## Introduction

A CRITICAL phase in any space mission is placing the spacecraft into the proper Earth orbit or escape trajectory. Obviously, a failure of the launch vehicle has catastrophic consequences for the mission. In the past, a number of failures of launch vehicles have occurred, notably the loss of the Space Shuttle *Challenger* at takeoff in 1986.[1] Assessing the risk of failure of launch vehicles can be difficult, especially for new vehicles that have had relatively few launch attempts. One approach to estimating the risk of launch vehicle failure is simply to use the actual past frequency of failure of that particular vehicle as an estimate of its failure probability. However, one can place little confidence in the results of this type of analysis unless there is a long history of launch attempts for the vehicle in question. When this is the case, two other approaches can be considered. The first is to use Bayesian probability (see Refs. 2–6). This approach relies on updating a prior probability distribution of the future frequency of launch successes based on the realized success rate of similar past vehicles. This Bayesian approach involves directly the launch vehicle as a whole without consideration of the performance of its subsystems. Bayesian updating has been used in a number of cases, for example, updating the probability that a nuclear attack is underway given evidence from a warning system[7] or computing the chance of a failure of the space shuttle based on past near misses.[8]

Another approach to the assessment of the likelihood of a launch vehicle failure in a given launch is to use probabilistic risk analysis (PRA). This requires decomposition of the launch vehicle into subsystems and components, assessment of the probability of failure of each component or subsystem, then computation of the probability of failure of the whole system based on fault tree and event tree analyses. The application of this approach and associated tools in the aerospace field have been well developed in the literature.[9,10]

In this paper, however, we focus on Bayesian methods and updating based on vehicle-level performance data rather than on PRA. We assess the probability that a given launch vehicle fails in a single launch, and we display the results of this analysis for the major launch vehicles in the world. We provide three different levels of analysis, each more sophisticated in the definition of the prior distribution of the failure frequency than the preceding levels. We introduce first the necessary background on Bayesian probability analysis. We then present these three levels of analysis.

## Bayesian Analysis Framework

The use of Bayesian probability in estimating system reliability is becoming more popular.[5,6,8] One of the main reasons is that it yields not only an estimate of system reliability but also a better estimate of the uncertainty about that estimate than classical statistical methods do. A second reason for this growing popularity is that Bayesian methods allow one to obtain an estimate of a system's reliability before the accrual of a large number of trials in operations.

Consider as a random variable the future frequency, $A$, of successful launches assuming that the system remains unchanged. Our objective here is to characterize the probability density function of $A$ given new experience in successive launches. Bayesian computation of probability distributions starts with a prior probability density function representing one's degree of belief about the different possible values for a parameter, then updates this distribution based on the observed data. This process relies on Baye's theorem:

$$f_{A|D}(a|s,t) = \frac{f_{D|A}(s,t|a)\,f_A(a)}{\int_x f_{D|A}(s,t|x)\,f_A(x)\,\mathrm{d}x} = \frac{(\text{likelihood}) \times (\text{prior})}{(\text{total probability of the data})} \tag{1}$$

In Eq. (1), $A$ is the future frequency of launch success given $s$ successful launches in $t$ past trials. Information about $A$ changes with experience. Consider a probability density function $f_A(a)$ that represents the previous state of knowledge about a given launch vehicle before any of the $t$ launch attempts are observed. This density

Received 16 January 2003; revision received 9 June 2003; accepted for publication 12 June 2003. 

*Ph.D. Student, Department of Management Science and Engineering. Student Member AIAA.

†Burt and Anne McMurty Professor, School of Engineering, and Chair, Department of Management Science and Engineering.

function is called a prior distribution. For example, using a uniform probability density function as a prior would imply that the expected future fraction of successful launches for a given vehicle is equally likely to lie anywhere between 0 and 1. The prior distribution can be based on expert assessment (e.g., Refs. 11 and 12), or, as we show further, from performance data of similar systems.

The likelihood function $f_{D|A}(s, t|a)$ represents the probability of the data $D$, that is, of obtaining $s$ successes in $t$ trials, given that the expected future fraction of launches successful is $a$. Again, different probability density functions can be used, depending on the characteristics of the situation. For example, the binomial distribution, which is discussed in detail later, is an appropriate likelihood function if the successive launches of a rocket can be viewed as a series of independent realizations of a Bernoulli random variable with a constant probability of success. The integral in the denominator of Eq. (1), called the preposterior, is a renormalization factor that makes the resulting distribution a valid probability density function. It is the probability of realizing the actual data for all possible future frequencies of events. The result of the calculation, $f_{A|D}(a|s, t)$, is the posterior density function that is obtained by updating the prior beliefs to reflect the realized data.

Howard[13] shows the general form of the updating process and proves that the best estimate of the probability of success in the next realization of the process is the mean of the posterior probability density function. He also shows that certain pairings of likelihood distributions with prior distributions, known as conjugate priors, simplify the calculations because the posterior distribution, the updated estimate of the future frequency of success, and prior distribution are of the same family of density functions. This permits simple algebraic computation of the posterior distribution rather than the use of numerical integration or simulation.

In the first two levels of our analysis of launch vehicle reliability, we use the conjugate prior distribution for a binomial likelihood function, which is a beta distribution. The formula for a beta probability density function is given by Eq. (2) with parameters $s$ and $f$, where $f = t - s$ for $0 \le a \le 1$ (Ref. 14). $\Gamma(n)$ is the gamma function defined as

$$\Gamma(n) = \int_0^\infty t^{n-1} e^{-t}\, dt$$

Thus,

$$f_A(a|s, t) = \frac{\Gamma(s + f)}{\Gamma(s)\Gamma(f)} a^{s-1}(1 - a)^{f-1} \tag{2}$$

The mean and variance of the beta distribution with parameters $s$ and $f$ are

$$\text{mean} = s/(s + f), \qquad \text{variance} = sf/[(s + f)^2(s + f + 1)] \tag{3}$$

As stated earlier, the beta distribution is the conjugate prior for the binomial likelihood function, which is given by the following equation for $0 \le a \le 1$ and $s < t$ (Ref. 14):

$$P_{s,t}(a; s, t) = \binom{t}{s} a^s (1 - a)^{t-s} \tag{4}$$

As shown in Fig. 1, the beta distribution can be used to model a wide range of different prior beliefs about $A$. For example, a uniform prior on the (0, 1) interval, which is also a beta (1, 1) distribution, represents a situation in which one believes that the future frequency of success is evenly distributed across the entire range (0, 1). This uninformative prior is often used to represent the absence of information, here about a system's robustness. A beta (0.5, 0.5) prior is similar to a beta (1, 1) prior except that it is heavier in the upper and lower tails. The beta (10, 10) prior could be used if one believes that the actual future frequency of success of a system is symmetrically distributed around a peak of 0.5. The beta (30, 5) prior could be used if one believed that the actual future frequency of success has a peak of 0.86. A beta (30, 5) prior distribution, thus, represents a prior belief that the system is relatively robust.

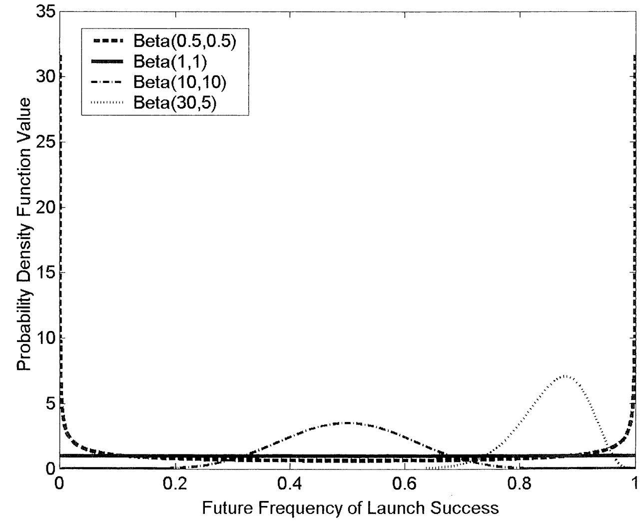



**Fig. 1 Examples of beta distributions for modeling the prior distribution of launch vehicle reliability.**

As shown by Howard,[13] the posterior distribution, the distribution for the future fraction of launches successful that takes into account both the prior information given by a beta $(s, f)$ distribution and the realized data ($s'$ successes in $t'$ launches), is given with $0 \le a \le 1$, and $0 < s' + s < t' + t$:

$$f_{A|D}(a|s', t') = \frac{\Gamma(s + f + t')}{\Gamma(s + s')\Gamma[f + (t' - s')]} a^{(s + s' - 1)} \times (1 - a)^{[f + (t' - s') - 1]} \tag{5}$$

In other words, if the prior is a beta distribution with parameters $s$ and $f$, and if the (new) data involve $s'$ successes in $t'$ trials that can be modeled as a Bernoulli process, the posterior distribution is also a beta distribution with parameters $s' + s$ and $f' + f$. This process is used in the first two levels of our analysis, updating a beta prior distribution with the available data. The third level uses a nonconjugate prior so that the Bayesian updating is done numerically. We next demonstrate this process by computing posterior probability density functions for the major families of launch vehicles in the world.

## Bayesian Analysis for Launch Vehicle Reliability

We performed a Bayesian analysis of the future frequency of success of the 33 major families of launch vehicles currently in use. By family, we mean the variants of a distinct launch vehicle from a particular manufacturer. For example, all models of the Delta 2 were included as one family. We obtained the data about the launch history of each vehicle from Isakowitz et al.[15] for launches before 30 June 1999 and from the launch logs compiled at Space.com[‡] for all launches between 30 June 1999 and the end of 2001. Any launch resulting in delivery of the payload to an orbit other than the intended orbit was considered a failure. Although the data on past launch failures could have been partitioned into failures in boost stages and failures in upper stages, they are aggregated in this paper. The data needed to classify all of the failures according to the stage in which they occurred were not available. The failure data and the results refer here to the launch vehicle systems as a whole, not to boost stages or upper stages individually. The data about past launches used in our analysis are shown in Table 1.

### First-Level Bayesian Analysis

In our first level of Bayesian analysis, we updated a uniform, or beta (1, 1), prior distribution with the launch data shown in Table 1

[‡]Space.com maintains lists of launches compiled by the newspaper *Florida Today* at http://www.space.com. We used the launch logs from 1999 to 2001, and we last accessed them 11 March 2002.

**Table 1 Launch data for different classes of launch vehicles**

| Launch vehicle family | Launch vehicle number | Number of successful launches | Number of launch failures | Realized success rate, % |
|---|---|---|---|---|
| Ariane 4 | 1 | 103 | 3 | 97.2 |
| Ariane 5 | 2 | 7 | 3 | 70.0 |
| Athena 1 and 2 | 3 | 5 | 2 | 71.4 |
| Atlas 2A and 2AS | 4 | 54 | 0 | 100.0 |
| Delta 2 | 5 | 94 | 2 | 97.9 |
| Delta 3 | 6 | 1 | 2 | 33.3 |
| Dnepr | 7 | 2 | 0 | 100.0 |
| GSLV (India) | 8 | 1 | 0 | 100.0 |
| H-2A (Japan) | 9 | 6 | 2 | 75.0 |
| Kosmos | 10 | 400 | 23 | 94.6 |
| Long March 2C | 11 | 22 | 0 | 100.0 |
| Long March 2E, 2F | 12 | 4 | 3 | 57.1 |
| Long March 3A, 3B, 3C | 13 | 19 | 5 | 79.2 |
| Long March 4, 4B, 2D | 14 | 8 | 0 | 100.0 |
| Minotaur | 15 | 2 | 0 | 100.0 |
| Molniya | 16 | 266 | 19 | 93.3 |
| M-V (Japan) | 17 | 2 | 1 | 66.7 |
| Pegasus XL | 18 | 18 | 3 | 85.7 |
| Proton K/Block M | 19 | 251 | 33 | 88.4 |
| PSLV (India) | 20 | 4 | 2 | 66.7 |
| Rockot | 21 | 2 | 0 | 100.0 |
| Shavit (Israel) | 22 | 3 | 1 | 75.0 |
| Soyuz | 23 | 671 | 20 | 97.1 |
| Space Shuttle | 24 | 105 | 1 | 99.1 |
| Start-1 | 25 | 5 | 1 | 83.3 |
| Strela (Russia) | 26 | 1 | 0 | 100.0 |
| Taurus | 27 | 5 | 1 | 83.3 |
| Titan 2 | 28 | 10 | 0 | 100.0 |
| Titan 4 | 29 | 10 | 1 | 90.9 |
| Tsyklon 2 and 3 | 30 | 230 | 8 | 96.6 |
| VLS-1 (Brazil) | 31 | 0 | 1 | 0.0 |
| Zenit 2 | 32 | 28 | 5 | 84.8 |
| Zenit 3 (sea launch) | 33 | 6 | 1 | 85.7 |
| All launch vehicles combined | | 2345 | 143 | 94.3 |

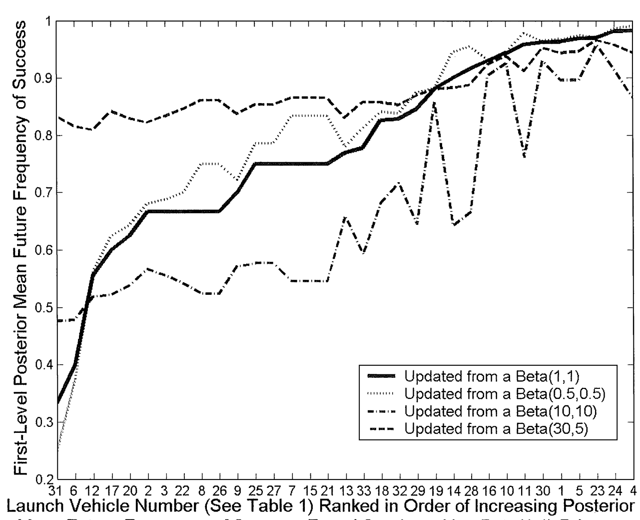



**Fig. 2 Comparison of first-level analysis with different prior distributions.**

through Eq. (5), and the means of the resulting beta posterior distributions are shown in Table 2. Figure 2 shows a comparison of the results of starting with different prior distributions by showing the posterior estimates of the mean future frequency of success starting with the four different beta prior distributions shown in Fig. 1. In Fig. 2, the launch vehicles have been presented in order of increasing posterior means (future frequency of success), based on a beta (1, 1) prior distribution on the $x$ axis. Points below this solid

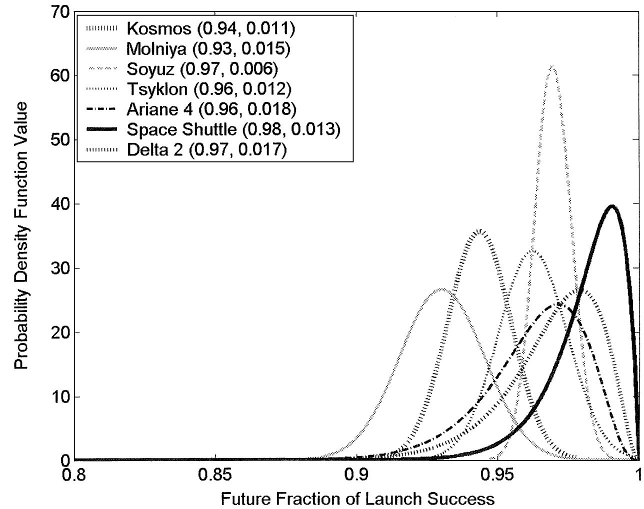



**Fig. 3 First-level posterior probability density functions for vehicles with at least 96 launch attempts; scale on the $x$ axis runs from 0.8 to 1.0. Mean and standard deviation for each distribution in legend are in parentheses.**

line represent posterior estimates that were lower than those based on a beta (1, 1) prior, and points above this line represent posterior estimates that were above those based on the uniform prior. Figure 2 shows that similar results are obtained by updating the beta (1, 1) and the beta (0.5, 0.5) distributions. However, for the vehicles with low mean rates of success, the differences between the posteriors estimated by updating the beta (10, 10) and beta (30, 5) distributions were quite large. These large differences occur because these distributions represent strong prior beliefs about the future frequency of success of the vehicles relative to the number of launch attempts of these vehicles. We use the beta (1, 1) as the first-level prior throughout the remainder of this paper.

Figure 3 shows the posterior density functions for the future frequency of launch success (the random variable $A$) along with the means and standard deviations of these distributions for vehicles with at least 96 launch attempts. Figure 4 shows the posterior density functions, means, and standard deviations for all rockets with at least 10 but fewer than 96 launch attempts, and Fig. 5 shows the posterior density functions, means, and standard deviations for rockets with fewer than 10 launch attempts. Again, as shown by Howard,[13] the best estimate of the probability of success in the next launch for a given launch vehicle is the mean of the posterior distribution for future frequency of launch success. This implies that, in a probabilistic risk analysis for a mission, it is the mean of the posterior that should be used when calculating the probability of mission failure due to a launch vehicle failure.

Figure 3 shows that of the launch vehicles with at least 96 launches, the space shuttle has the highest posterior mean future frequency of success, whereas the Soyuz, the rocket that has the highest number of launch attempts, has the narrowest posterior density function, that is, the lowest standard deviation.

As compared to the launch vehicles shown in Fig. 3, the vehicles shown in Fig. 4 generally have higher variances about the means due to the smaller number of launch attempts. Of the vehicles included in Fig. 4, the Atlas 2 and Long March 2C have the highest probability of success, and the Atlas 2 has the lowest standard deviation.

The rockets graphed in Fig. 5 again have higher variances than those in Fig. 4 because of the lower numbers of launch attempts. The means also vary considerably from 0.33 for the VLS-1, a Brazilian launch vehicle, to 0.90 for the Long March 4 and 2D family of vehicles.

Using a uniform beta (1, 1) prior assumes that we know nothing about the reliability of launch vehicles in general before the launch of a given vehicle. In reality, we know a considerable amount about launch vehicles from the many different types currently in use. As shown in Table 1, there have been over 2300 launch attempts in the past for the launch vehicle families that are still in use. The next step

**Table 2 Means of the posterior distributions of future success rates for the three levels of analysis**

| | | | Posterior estimates of the mean future frequency of success | | | |
|---|---|---|---|---|---|---|
| Launch vehicle | Launch successes/ Launch attempts | Realized success rate | First level | Second-level interpolation fit | Second-level method of moments fit | Third level |
| Ariane 4 | 103/106 | 0.97 | 0.96 | 0.97 | 0.96 | 0.97 |
| Ariane 5 | 7/10 | 0.70 | 0.67 | 0.71 | 0.73 | 0.72 |
| Athena 1 and 2 | 5/7 | 0.71 | 0.67 | 0.73 | 0.74 | 0.74 |
| Atlas 2A and 2AS | 54/54 | 1.0 | 0.98 | 0.99 | 0.98 | 0.98 |
| Delta 2 | 94/96 | 0.98 | 0.97 | 0.97 | 0.97 | 0.97 |
| **Delta 3** | **1/3** | **0.33** | **0.40** | **0.63** | **0.62** | **0.57** |
| **Dnepr** | **2/2** | **1.0** | **0.75** | **0.85** | **0.84** | **0.87** |
| **GSLV (India)** | **1/1** | **1.0** | **0.67** | **0.82** | **0.81** | **0.84** |
| H-2A (Japan) | 6/8 | 0.75 | 0.70 | 0.74 | 0.76 | 0.76 |
| Kosmos | 400/423 | 0.95 | 0.94 | 0.94 | 0.94 | 0.95 |
| Long March 2C | 22/22 | 1.0 | 0.96 | 0.97 | 0.96 | 0.97 |
| Long March 2E, 2F | 4/7 | 0.57 | 0.56 | 0.67 | 0.66 | 0.64 |
| Long March 3A, 3B, 3C | 19/24 | 0.79 | 0.77 | 0.76 | 0.79 | 0.79 |
| Long March 4, 4B, 2D | 8/8 | 1.0 | 0.90 | 0.93 | 0.91 | 0.94 |
| **Minotaur** | **2/2** | **1.0** | **0.75** | **0.85** | **0.84** | **0.87** |
| Molniya | 266/285 | 0.93 | 0.93 | 0.93 | 0.93 | 0.93 |
| M-V (Japan) | 2/3 | 0.67 | 0.60 | 0.73 | 0.74 | 0.74 |
| Pegasus XL | 18/21 | 0.86 | 0.83 | 0.83 | 0.84 | 0.85 |
| Proton K/Block M | 251/284 | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 |
| PSLV (India) | 4/6 | 0.67 | 0.63 | 0.71 | 0.72 | 0.71 |
| **Rockot** | **2/2** | **1.0** | **0.75** | **0.85** | **0.84** | **0.87** |
| Shavit (Israel) | 3/4 | 0.75 | 0.67 | 0.75 | 0.77 | 0.77 |
| Soyuz | 671/691 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 |
| Space shuttle | 105/106 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 |
| Start-1 | 5/6 | 0.83 | 0.75 | 0.79 | 0.81 | 0.82 |
| **Strela (Russia)** | **1/1** | **1.0** | **0.67** | **0.82** | **0.81** | **0.84** |
| Taurus | 5/6 | 0.83 | 0.75 | 0.79 | 0.81 | 0.82 |
| Titan 2 | 10/10 | 1.0 | 0.92 | 0.94 | 0.92 | 0.94 |
| Titan 4 | 10/11 | 0.91 | 0.85 | 0.87 | 0.87 | 0.89 |
| Tsyklon 2 and 3 | 230/238 | 0.97 | 0.96 | 0.96 | 0.96 | 0.96 |
| **VLS-1 (Brazil)** | **0/1** | **0.0** | **0.33** | **0.66** | **0.65** | **0.61** |
| Zenit 2 | 28/33 | 0.85 | 0.83 | 0.82 | 0.84 | 0.84 |
| Zenit 3 (sea launch) | 6/7 | 0.86 | 0.78 | 0.81 | 0.82 | 0.84 |

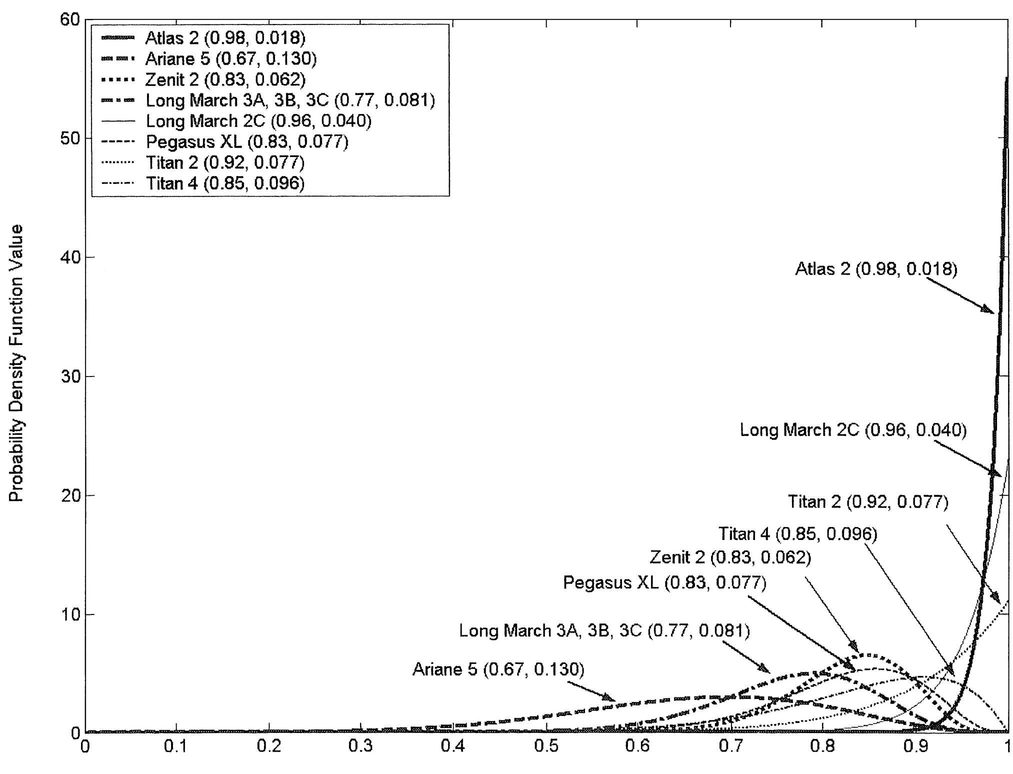



**Fig. 4 First-level posterior probability density functions for vehicles with more than 10 but fewer than 96 launch attempts. Mean and standard deviation for each distribution in legend are in parentheses.**

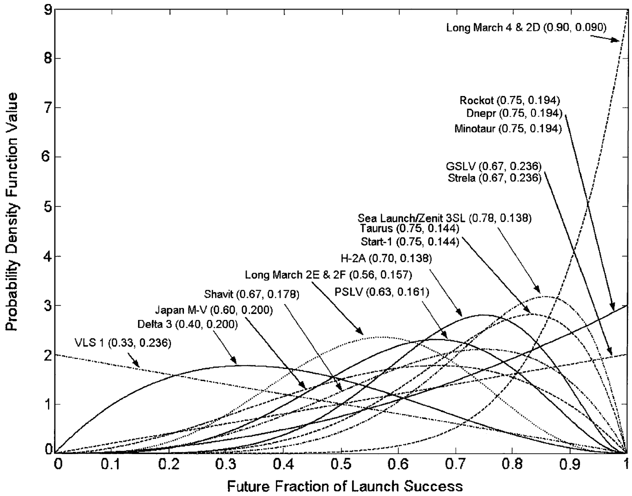



**Fig. 5 First-level posterior probability density functions for vehicles with fewer than 10 launch attempts. Mean and standard deviation for each distribution in legend are in parentheses.**

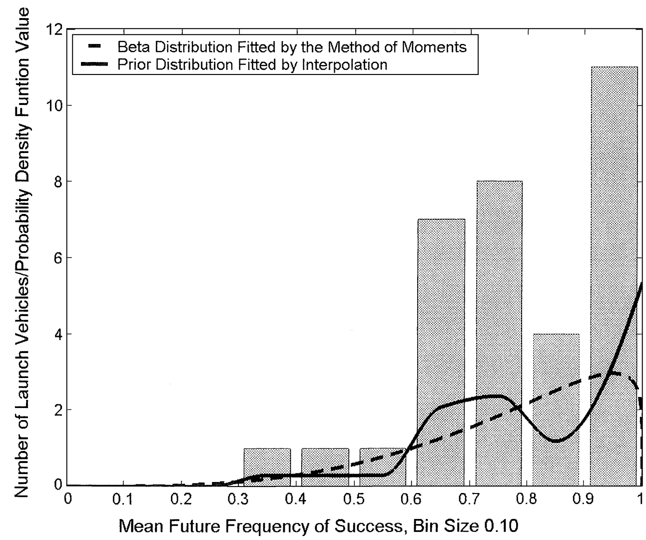



**Fig. 6 Histogram of first-level posterior means and fitted second-level prior distributions.**

in our analysis uses this information to improve the formulation of the prior density function.

### Second-Level Bayesian Analysis

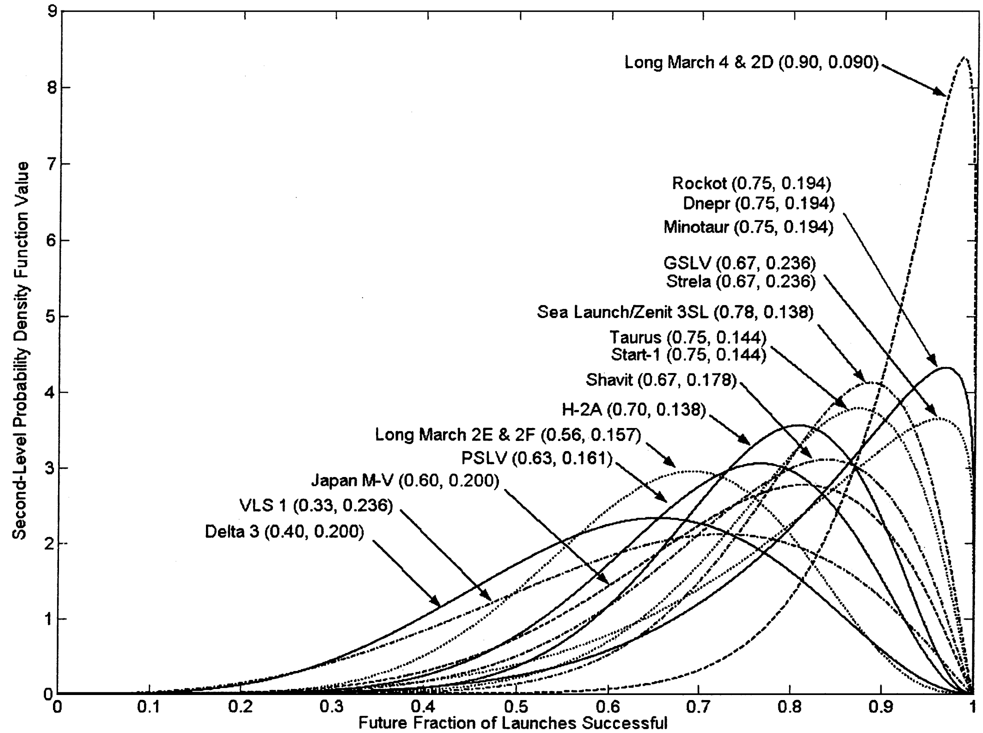
We now incorporate the data about past launches into the second-level prior distribution by using the means of the first-level posterior distributions to fit a new prior distribution. This approach implies that before observing the launch record of a new vehicle, one believes that this vehicle is no more and no less likely to fail in a launch attempt than an average existing vehicle. This approach might, thus, introduce some bias into the resulting estimates of the future frequency of launch success for those vehicles that are either better or worse, but we generally do not know. Because we weigh all launch vehicles equally, those with long launch histories have a larger impact on our posterior estimates than those with shorter launch histories. Vehicles with short launch histories are either new and have not been in use long enough for many launch attempts or are perceived as unreliable or costly relative to the other vehicles and are, thus, not selected often. By basing our second-level prior on the launch histories of all vehicles, equally weighted, we may be implying that a new vehicle is more similar to the vehicles with a long launch history, that is, those that are more likely to be robust, than those with fewer launches. This could introduce a positive bias (toward higher estimates of the mean future frequency of launch success) into our results. In the absence of additional information leading to a different prior belief about the success rate of a particular vehicle, the potential bias in our approach should not be a problem because our approach makes best use of the available data. If, however, available information suggests that our assumption about the prior is not reasonable, these data should be used to modify the prior. For example, in some cases, design changes have been made to improve a vehicle. As additional data from launches after these modifications are incorporated into the data set, the success rate estimated by our procedure would reflect the effects of these design changes. However, we have not included here a procedure to incorporate knowledge about design changes in the prior distribution (before some observations of the actual performance of the modified system). Other authors have developed methods to address this problem.[5]

In the second level of our analysis we use two different prior distributions. First, we fitted a beta distribution to the means of the first-level posteriors by the method of moments. Second, we fitted a curve to the histogram of the first-order posterior means by interpolation. The former (the method of moments) has the advantage of matching the mean and variance of the second-level prior distribution to the mean and variance of the means of the first-level posterior distributions, and it provides a conjugate prior. The latter (curve fitting by interpolation) has the advantage of "looking" most like the histogram of the first-level posterior means, but it requires numerical integration in the updating process. Figure 6 shows both the histogram of the first-level posterior means and the results of the two different approaches to fitting the second-level beta prior. The prior fitted by interpolation has a mean of approximately 0.79, whereas the prior fitted by the method of moments, a beta (4.10, 1.16) distribution, has a mean of approximately 0.78.

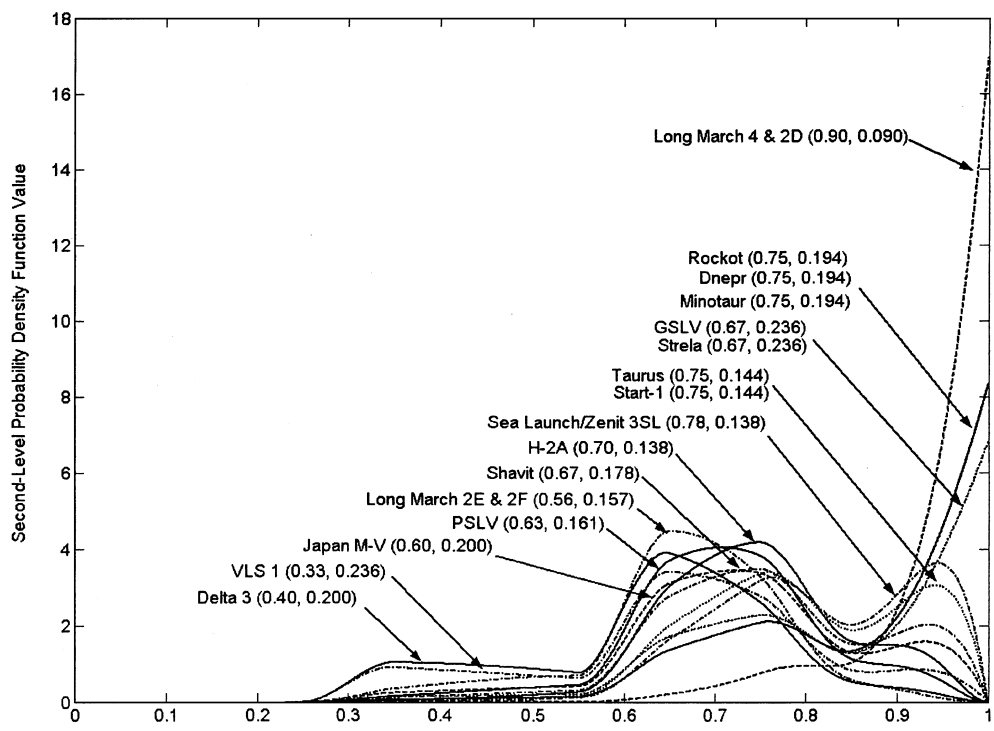
The choice between the two different ways to fit these distributions is a judgment call. We present the results from both, and we examine the differences in the final results. We update these second-level priors with the launch data for each vehicle using Eq. (5) for the method of moments fit and numerical integration for the interpolation fit. As an example of the results of the second-level analysis, Fig. 7 shows the posterior density functions, means, and standard deviations for rockets with fewer than 10 launch attempts. Figure 7 shows that, for the launch vehicles with no more than 10 launch attempts, the posterior distributions are very different when starting from the two different prior distributions. These differences are to be expected because the prior distributions matter a great deal for vehicles with few launch data but little for vehicles with a large amount of data. Although not shown here, the posterior density functions for vehicles with more than 10 launch attempts are similar, except that the differences between the results starting from the two different priors do not differ as much for vehicles with larger numbers of past launches. The full results of the second-level analysis are summarized in Tables 2 and 3.

### Third-Level Bayesian Analysis

To estimate better the probability of success of a launch vehicle, our third-level Bayesian analysis uses the first-level posterior distributions for all of the launch vehicles by combining the first-level posteriors into a new nonbeta distribution by summing all of the first-level posteriors and renormalizing the resulting function into a proper density function. This new prior distribution then represents the best estimate of the future frequency of launch success for a new vehicle before the results of any launch attempt are known, assuming that the new vehicle is similar in its reliability to those currently in use. One way of testing this assumption would be to compare the results for new launch vehicle operations in the future with the third-order prior developed here. We present such a comparison in a subsequent section. Figure 8 shows the new third-level prior distribution computed by combining and renormalizing the first-level posterior distributions. The mean and variance of this third-level prior distribution are approximately 0.78 and 0.044, respectively. This mean would then be the best estimate of the probability of success of a new launch vehicle, such as the Delta 4, before




**a) Second-level posterior distributions with the method of moments fit for the prior**




**b) Second-level posterior distributions with interpolation fit for the prior**

**Fig. 7 Second-level posterior probability density functions for vehicles with fewer than 10 launch attempts. Mean and standard deviation for each distribution in legend are in parentheses.**

observing the results of launch attempts, assuming that the new vehicle is no more and no less reliable than the existing launch vehicles as a group.

To update the probability of success of a launch vehicle after a number of launch attempts have been made, the third-level prior is updated with the launch data. Because the third-level prior is no longer a conjugate prior for the binomial likelihood function, the updating must be done numerically. We use trapezoid integration with 10,000 points over the interval (0, 1) in this updating; for example, see, Ref. 16. We selected 10,000 points because for a lower number of points, for example, 1000, the results were still sensitive to changes in the number of points used, but with 10,000 points the results did not change for small variations of the number of points in the interval. For launch vehicles that already have a performance record, we compute the third-level prior distribution by eliminating the first-level posterior of the launch vehicle for which a third-level

**Table 3 Standard deviations of the posterior distributions of future success rates for the Bayesian analysis**

| Launch vehicle | First level | Second-level interpolation fit | Second-level method of moments fit | Third level |
|---|---|---|---|---|
| Ariane 4 | 0.0181 | 0.0164 | 0.0179 | 0.0158 |
| Ariane 5 | 0.1307 | 0.0933 | 0.1104 | 0.1197 |
| Athena 1 and 2 | 0.1491 | 0.1063 | 0.1201 | 0.1329 |
| Atlas 2A and 2AS | 0.0175 | 0.0539 | 0.0179 | 0.0306 |
| Delta 2 | 0.0173 | 0.0156 | 0.0172 | 0.0152 |
| Delta 3 | 0.2000 | 0.1431 | 0.1597 | 0.1888 |
| Dnepr | 0.1936 | 0.1334 | 0.1276 | 0.1331 |
| GSLV (India) | 0.2357 | 0.1466 | 0.1443 | 0.1604 |
| H-2A (Japan) | 0.1382 | 0.1032 | 0.1128 | 0.1240 |
| Kosmos | 0.0112 | 0.0109 | 0.0111 | 0.0114 |
| Long March 2C | 0.0400 | 0.0238 | 0.0380 | 0.0144 |
| Long March 2E, 2F | 0.1571 | 0.1121 | 0.1300 | 0.1439 |
| Long March 3A, 3B, 3C | 0.0811 | 0.0735 | 0.0741 | 0.0782 |
| Long March 4, 4B, 2D | 0.0905 | 0.0745 | 0.0749 | 0.0634 |
| Minotaur | 0.1936 | 0.1334 | 0.1276 | 0.1331 |
| Molniya | 0.0150 | 0.0144 | 0.0149 | 0.0152 |
| M-V (Japan) | 0.2000 | 0.1331 | 0.1444 | 0.1678 |
| Pegasus XL | 0.0774 | 0.0829 | 0.0699 | 0.0745 |
| Proton K/Block M | 0.0191 | 0.0194 | 0.0189 | 0.0198 |
| PSLV (India) | 0.1614 | 0.1115 | 0.1283 | 0.1433 |
| Rockot | 0.1936 | 0.1334 | 0.1276 | 0.1331 |
| Shavit (Israel) | 0.1782 | 0.1258 | 0.1320 | 0.1495 |
| Soyuz | 0.0065 | 0.0064 | 0.0065 | 0.0064 |
| Space Shuttle | 0.0129 | 0.0118 | 0.0130 | 0.0125 |
| Start-1 | 0.1443 | 0.1167 | 0.1125 | 0.1228 |
| Strela (Russia) | 0.2357 | 0.1466 | 0.1443 | 0.1604 |
| Taurus | 0.1443 | 0.1167 | 0.1125 | 0.1228 |
| Titan 2 | 0.0767 | 0.0597 | 0.0658 | 0.0526 |
| Titan 4 | 0.0964 | 0.0967 | 0.0817 | 0.0835 |
| Tsyklon 2 and 3 | 0.0122 | 0.0117 | 0.0122 | 0.0115 |
| VLS-1 (Brazil) | 0.2357 | 0.1562 | 0.1764 | 0.2177 |
| Zenit 2 | 0.0628 | 0.0693 | 0.0587 | 0.0620 |
| Zenit 3 (sea launch) | 0.1315 | 0.1131 | 0.1047 | 0.1125 |

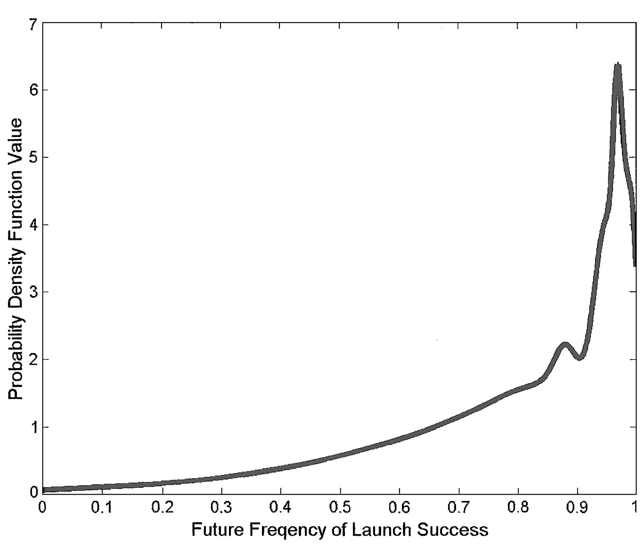



**Fig. 8 Third-level prior probability density functions.**

posterior estimate is being sought from the prior. That is, we recompute the third-level prior by using all of the first-level posteriors except the one for the vehicle under consideration. We then numerically update this prior with the available launch data for the vehicle in question. As examples of the results of the third-level analysis, Fig. 9 shows the third-level posterior density functions, means, and standard deviations for vehicles with at least 96 launch attempts, and Fig. 10 shows the third-level posterior density functions, means, and standard deviations for all rockets with fewer than 10 launch attempts. The third-level posteriors with vehicles with between 10 and 96 launch attempts are not shown here but are summarized in Tables 2 and 3.

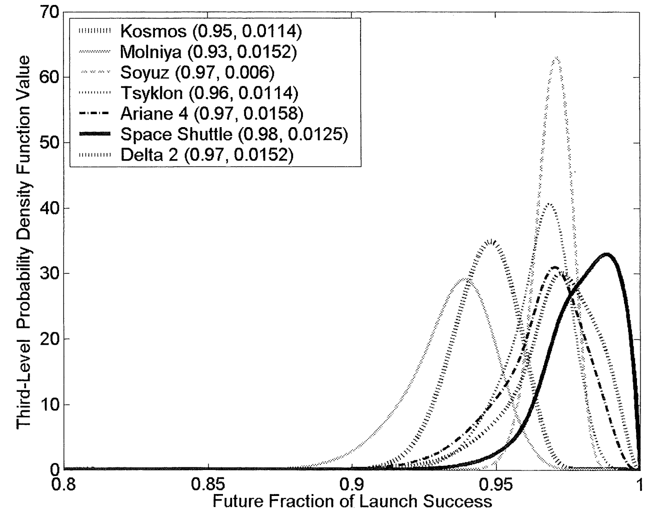



**Fig. 9 Third-level posterior probability density functions for vehicles with at least 96 launch attempts. Mean and standard deviation for each distribution in legend are in parentheses.**

Comparing the results of the three levels of Bayesian analysis (Table 2), we see that, for the vehicles with at least 96 launch attempts, the posterior means from the three different levels of Bayesian analysis differ by less than 2%. Furthermore, as is shown in Table 2, the posterior mean estimates for vehicles with at least 96 launch attempts generally differ little from the classical mean computed by dividing the number of successful launches by the total number of launch attempts. For these rockets, there are enough data so that the prior distribution has little influence on the posterior because the data overwhelm the effect of the prior. However, even if the mean success rate estimates do not differ greatly between the different methods, the third-level Bayesian approach does give a better understanding of the uncertainties involved. In some cases this method will lead to higher uncertainty estimates, whereas in others the uncertainty estimates will be lower. As always, no general statement can be made that the use of Bayesian methods reduces uncertainty because in some cases making better use of the data in computing the system's success rates increases the estimates of the uncertainty about the success rate estimates.

As shown in Table 2, the posterior means computed by the three methods differ by less than 5% for most of the vehicles with 10–95 launches. The three exceptions are the Ariane 5, the Long March 3, and the Titan 4. Figure 7 also shows that, when nonconjugate prior distributions are used, the posterior density functions may be multimodal. However, the posterior means are still the best estimates of the future success rate for a given launch vehicle, given our assumptions.

Table 2 shows that for vehicles with a small amount of launch data, the prior distribution chosen for the analysis has a large impact on the resulting estimates of the future frequency of launch success. Both the shapes of the posterior density functions and the means change considerably for the vehicles shown in Figs. 5 and 10, depending on the prior distribution used in the analysis, and it is for vehicles with few launches that Bayesian analysis is most useful. Figure 10 also shows that many of the third-level posterior density functions are bimodal due to the absence of conjugate priors.

### Bayesian Analysis of First Launches

We now turn to the infancy problem, the possibility that a launch vehicle may be more likely to fail on its first launch attempt than later. We did this by collecting data from the first launches of 35 vehicles, then by updating a uniform [beta (1, 1)] prior. Our data set contained 28 successes in 35 first launch attempts, and the posterior probability density function for the mean future frequency of success on the first launch of a vehicle was a beta (29, 8) as shown in Fig. 11.

The mean and standard deviation of the infancy failure rate posterior distribution shown in Fig. 11 are 0.784 and 0.004, respectively.

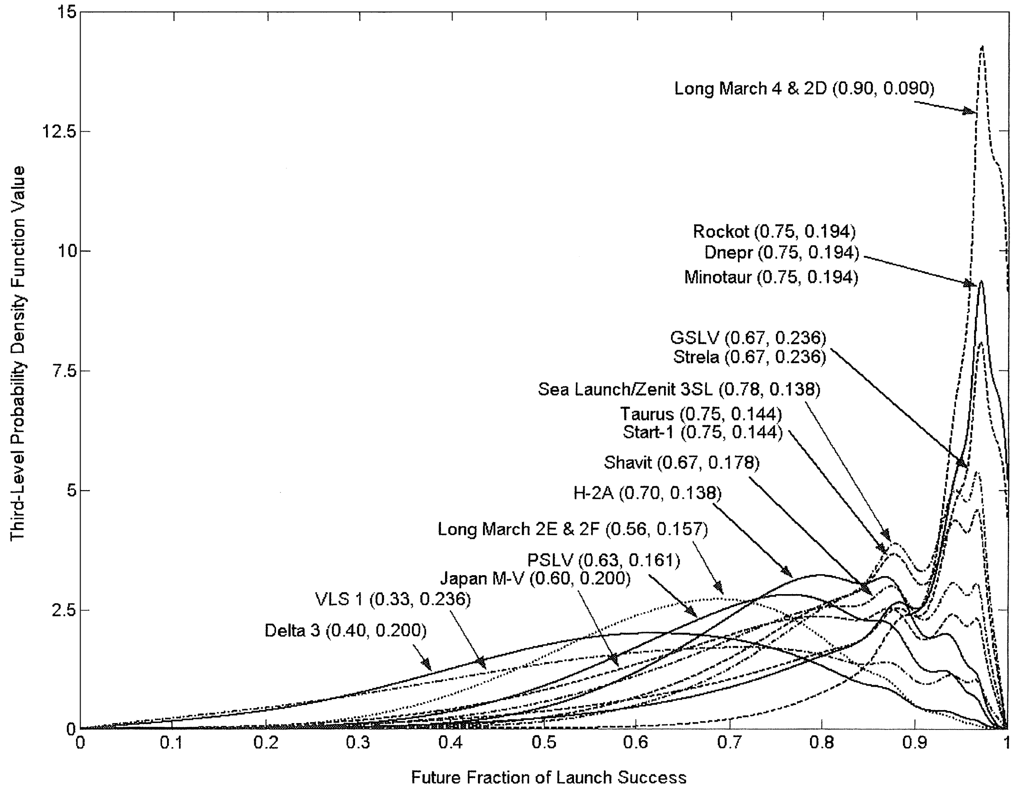



**Fig. 10 Third-level posterior probability density functions for vehicles with fewer than 10 launch attempts. Mean and standard deviation for each distribution in legend are in parentheses.**

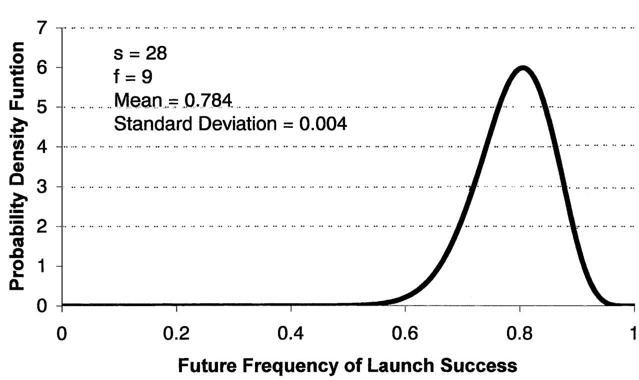



**Fig. 11 Posterior probability density function for the infancy problem; updated from uniform prior distribution.**

Comparing these statistics with the mean and variance for the computed third-level prior density function shown in Fig. 8, we see that the mean is nearly the same (0.78 vs 0.784), whereas the standard deviation of the infancy failure rate posterior distribution is an order of magnitude lower (0.044 vs 0.004). The mean future frequency of first launch success was, thus, found to be approximately the same by two different methods. This similarity of the first moment suggests that the results are probably good indicators of launch vehicle reliability in the infancy stage.

### Summary of the Bayesian Analyses

Table 2 summarizes the results of the three levels of the Bayesian analysis in terms of the mean future frequency of success of the posterior distributions found for each vehicle type by the different methods. Table 2 also includes the classical statistical mean (total number of successes divided by the total number of launches) for comparison.

A number of launch vehicles with large differences among the mean estimates of the future frequency of success obtained by different methods are highlighted with boldfaced type in Table 2. Note that, for all of these vehicles, we have relatively few launch data. It is precisely these types of vehicles for which Bayesian estimates of the future frequency of launch success are most useful because, by using Bayesian methods, one obtains not only an estimate of the mean, but the entire posterior probability density function of the success rate. For example, for the VLS-1 that failed in its only launch attempt in our data set (Table 1), using a numerical mean would suggest that its probability of success in the next launch was zero. However, using a Bayesian approach gives an estimated probability of success between 0.33 and 0.89, depending on what prior knowledge is assumed about the vehicle. Because the priors are important when there are few launch data, we use three different priors to examine the impact of the priors. This is the approach for Bayesian analysis that was recommended by Gelman et al.[17]

In contrast to the VLS-1, the estimates of the probability of success of a Soyuz on its next launch are essentially the same for the five different methods: 0.97. This is because the Soyuz has been launched many times (691 times as of the end of 2001, Table 1). Therefore, the available data overwhelm any reasonable prior probability distribution. Table 3 shows the standard deviations of the posterior distributions found with each of the three levels of Bayesian analysis for each launch vehicle.

We next examined the question of which, if any, of the methods provide different estimates of the probability of success for the different rockets. To answer this question we used pairwise hypothesis testing using classical (frequentist) statistics to compare the results of the various reliability estimation procedures. We began by computing the average (over all launch vehicles) of the posterior mean success rates found by the four versions of Bayesian analysis and by the statistical fraction of successful launches. We then performed pairwise hypothesis test comparisons of these five averages and found that only the estimates found through the first level of Bayesian analysis were significantly different from the classical statistical mean ($p$ value $= 0.10$). However, when performing hypothesis tests to compare the predicted means obtained by classical statistical analysis, that is, the number of successes divided by the number of launch attempts, and by the three levels of Bayesian analysis, the $p$ values were 0.05, 0.04, and 0.04, respectively, for the two second-level approaches and the third-level approach in comparison to the first level. All of these pairwise hypothesis tests were done by comparing the null hypothesis of equal averages (of the posterior rates of success over all vehicles) with the alternative hypothesis that the second method in the test yielded a higher posterior rate of success on average using a two-sample $t$ test.[14] The results, thus, suggest that the second and third levels of Bayesian methods yield, on average, estimates of the posterior mean future frequency of success that are similar to those of the classical statistical approach (fraction of successful launches). In this respect, one may think that the second and third levels of Bayesian analysis are more objective as conceptualized by Vaurio[12] because they are consistent with the classical approach. However, unlike the classical statistical method, the Bayesian methods provide not only a point estimate of the probability of success (the mean of the future frequency of success) but also the full probability density function.

## Comparison of Classical and Bayesian Confidence Limits

To compare the classical approach of computing the statistical mean and a confidence interval around that mean with the Bayesian approach, the 95% confidence interval for the proportion of successful launches, that is, the frequency of launch success, was computed using the classical formula[18]

$$p - z_{a/2}\sqrt{[p(1-p)]/n} \leq p \leq p + z_{a/2}\sqrt{[p(1-p)]/n} \qquad (6)$$

where $p$ is the estimate of the future frequency of success, $z_{a/2}$ is the $a/2$ significance level from the standard normal distribution, for example, 1.96 for a two-sided 95% confidence interval, and $n$ is the total number of launches. An approximate Bayesian 95% confidence interval was also computed by finding the 2.5th percentile and 97.5th percentile of the first-level posterior cumulative density functions for each of the launch vehicles. These two confidence interval estimates are shown in Fig. 12, together with the statistical mean based on the launch history of each vehicle and the first-level posterior estimate of the mean future frequency of success for each vehicle. The results show that the Bayesian approach provides a narrower confidence interval in nearly all cases. In addition, the Bayesian confidence interval encloses the classical mean for most vehicles in addition to the first-level Bayesian posterior mean. However, in some cases, for example, the VLS-1 (vehicle 31), the Bayesian confidence interval does not enclose the statistical mean. Because this vehicle has failed in its only launch attempt, the classical approach assigns a probability of zero to the future probability of success, which is clearly wrong. By contrast, the Bayesian method considers a larger experience base and, thus, assigns a smaller value (than 1) to the chance that the VLS-1 will never succeed in future launches, that is, that it will have a future frequency of success of 0 as predicted by the classical approach.

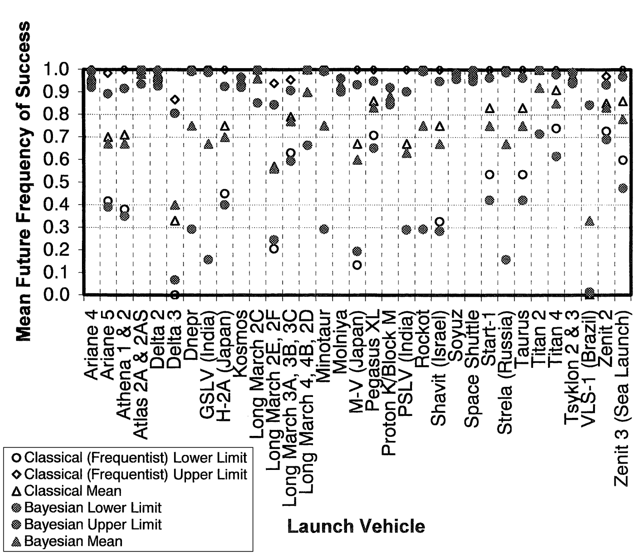



**Fig. 12 Classical (frequentist) and approximate Bayesian 95% confidence intervals.**

## Discussion

We have presented three levels of Bayesian analysis for estimating the future frequency of success for launch vehicles, and we have compared them to the classical statistical approach. The main benefits of the Bayesian approach are the following. First, it allows estimation of the future frequency of success of launch vehicles for which no launch attempts have been made. Second, it yields the entire probability density function for the future frequency of launch success instead of a point estimate. This is particularly important for vehicles with few previous launch attempts. Third, it provides tighter confidence intervals than a classical approach in many cases. Fourth, it can be done without decomposing the system into subsystems and components.

When the probability of success of a vehicle on the next launch is estimated, our first level of Bayesian analysis is appropriate if one can assume that nothing is known about launch vehicle reliability before observing the launch attempts of a given vehicle. However, there is a considerable amount of experience with launch vehicles, and this experience should be taken into account in the prior distribution as demonstrated in our analysis. Our second and third levels take the past experience into account in different ways. The second level considers only the past mean rates of success of the existing launch vehicles in defining the new prior distribution, whereas the third level incorporates all of the information provided by the full first-level posterior density functions.

The two main limitations of our analysis are that it does not take into account the chronological history of the different launch vehicles and that it does not account for similarities between new and past vehicles. The chronological history of launch vehicles is important because one would hope that, as a given launch vehicle matures, its probability of success increases as system problems are found and fixed before additional launch attempts. Therefore, the failure rate of a launch vehicle decreases as the infancy problems are resolved. We have assumed in our analysis that the estimated probability of success before the incorporation of additional data points is constant over the life of a launch vehicle, but we do update our estimate of this success rate as new data become available. Additional work could investigate methods for directly incorporating changes in the failure rate into the Bayesian analysis, possibly by weighting differently the various segments of the realized data.

Also, we have not directly taken into account the similarities between new vehicles and existing vehicles. Our method estimates the probability of failure of a new launch vehicle with no launch history by updating the third-level prior distribution shown in Fig. 8, a prior based on an equal weighting of all launch vehicles. This implicitly assumes that, before observing the results of launch attempts for the new launch vehicle, no information about its reliability relative to other vehicles is available. In some cases, this may not be true. For example, if the new vehicle represents only an incremental change with respect to an existing one, for example, the Delta 4 represents in many ways an incremental change from the previous Delta vehicles, it may share some of the reliability characteristics of the previous launch vehicles. Furthermore, there may be additional similarities between a new vehicle and previous ones due to the use of similar subsystems that may reduce the chance of some of the failure modes of past vehicles. However a new hybrid-fuel vehicle, that is, one using both liquid and solid propellants, would have a different set of potential failure modes than a solid propellant vehicle. Decomposing the data by failure mode and using this more detailed level of data in the Bayesian analysis for a new vehicle could provide better estimates of the future frequency of success.

Therefore, despite these two limitations, the Bayesian approach presented here to estimate the probability of success of the next launch attempt of a given vehicle has significant advantages over existing (classical statistics) methods. It allows estimation of the probability of failure of a launch vehicle and of the uncertainty about it before a large number of launch attempts have been made, without relying on a decomposition of the system through probabilistic risk analysis.

## Conclusions

We have presented a Bayesian method for estimating the future frequency of success of launch vehicles based on their past performance records. This method allows launch vehicle success rates to be estimated without a full decomposition of the system through PRA, and it overcomes some of the limitations of classical statistical analyses. The Bayesian approach developed here provides not only a point estimate of a launch vehicle's success (or failure) rate, but a full probability density function for that rate. Therefore, the Bayesian approach gives a much better description of the uncertainty about success rate estimates. It can also be used when there are relatively few data, making it particularly valuable for estimating the future success rates of new vehicles. The method presented can help support the selection of both a launch vehicle and a level of insurance coverage for a given mission.

## Acknowledgments

This research was partially funded by a National Defense Science and Engineering Graduate Fellowship from the U.S. Department of Defense, by Grant SES-0079872 from the National Science Foundation, and by the Space Systems, Policy, Architecture and Research Consortium grant to the Massachusetts Institute of Technology, California Institute of Technology and Stanford University from the U.S government.

## References

[1]Chang, I., "Investigation of Space Launch Vehicle Catastrophic Failures." *Journal of Spacecraft and Rockets*, Vol. 33, No. 2, 1996, pp. 198–205.

[2]Benjamin, J. R., and Cornell, C. A., *Probability, Statistics, and Decision for Civil Engineers*, McGraw–Hill, New York, 1970.

[3]Savage, L. J., *The Foundations of Statistics*, Wiley, New York, 1954.

[4]Raiffa, H., and Schlaifer, R., *Applied Statistical Decision Theory*, MIT Press, Boston, 1968.

[5]Gallagher, M. A., Weir, J. D., and True, W. D., "Relating Weapon System Test Sizes to Warfighting Capability," *Military Operations Research*, Vol. 3, No. 3, 1992, pp. 5–12.

[6]Fries, A., and Easterling, R. G., "Annual Testing of Strategic Missile Systems," *Proceedings of the American Statistical Association Section on Physical and Engineering Sciences*, American Statistical Association, Alexandria, VA, 1999, pp. 132–134.

[7]Paté-Cornell, M. E., and Fischbeck, P. S., "Probabilistic Interpretation of Command and Control Signals: Bayesian Updating of the Probability of Nuclear Attack," *Reliability Engineering and System Safety*, Vol. 47, No. 1, 1995, pp. 27–36.

[8]Kaplan, S., "On the Inclusion of Precursor and Near Miss Events in Qualitative Risk Assessments: A Bayesian Point of View and a Space Shuttle Example," *Reliability Engineering and System Safety*, Vol. 27, No. 1, 1990, pp. 103–115.

[9]Guarro, S., Bream, B., Rudolph, L. K., and Mulvihill, R. J., "The Cassini Mission Risk Assessment Framework and Application Techniques," *Reliability Engineering and System Safety*, Vol. 49, No. 3, 1995, pp. 293–302.

[10]Paté-Cornell, M. E., and Fischbeck, P. S., "PRA as a Management Tool: Organizational Factors and Risk-Based Priorities for the Maintenance of the Tiles of the Space Shuttle Orbiter," *Reliability Engineering and System Safety*, Vol. 40, No. 3, 1993, pp. 239–257.

[11]Spetzler, C. S., and von Holstein, C. A. S., "Probability Encoding in Decision Analysis," *Management Science*, Vol. 22, No. 3, 1975, pp. 340–354.

[12]Vaurio, J. K., "Objective Prior Distributions and Bayesian Updating." *Reliability Engineering and System Safety*, Vol. 35, No. 1, 1992, pp. 55–59.

[13]Howard, R. A., "Decision Analysis: Perspectives on Inference, Decision, and Experimentation," *Proceedings of the IEEE*, Vol. 58, No. 5, 1970, pp. 823–834.

[14]Devore, J. L., *Probability and Statistics for Engineering and the Sciences*, 3rd ed., Brooks/Cole, Pacific Grove, CA, 1991.

[15]Isakowitz, S. J., Hopkins, J. P., Jr., and Hopkins, J. B., *International Reference Guide to Space Launch Systems*, 3rd ed., AIAA, Reston, VA, 1999.

[16]Chapra, S. C., and Canale, R. P., *Numerical Methods for Engineers*, 2nd ed., McGraw–Hill, New York, 1988, pp. 478–489.

[17]Gelman, A. B., Carlin, J. S., Stern, H. S., and Rubin, D. B., *Bayesian Data Analysis*, Chapman and Hall, Boca Raton, FL, 2000, Chap. 6.

[18]Mitra, A., *Fundamentals of Quality Control and Improvement*, Prentice–Hall, Upper Saddle River, NJ, 1998, p. 111.

J. Korte
*Guest Editor*