Launch Vehicle Design and Requirements Verification Using Statistical Methods

November 8, 2023

Dr. John Hanson NASA/MSFC

Outline

- **This presentation applies statistics to launch vehicle design, but the methods may be used for other engineering applications**
- **Vehicle Models: What do we know when? How to separate parameters for making Monte Carlo runs**
  - Known when assembling the vehicle but not during design
  - Known prior to committing to flight
  - Unknown at lift off
  - How do we correctly model the vehicle during the design phases?
- **Number of Monte Carlo samples**
  - Requirements success
  - Design parameter values

Vehicle Models

Introduction: Varying Parameters

- Vehicle parameters are estimated during the design phases
  - Uncertainties are relatively large
- Many of these parameters will be known better for a specific rocket
  - We weigh the components for example
- Some things are better known on flight day
  - Estimate of the winds for example
  - Known parameters can be used to design the flight day trajectory
- There remains uncertainty on flight day

Varying Vehicle Parameters

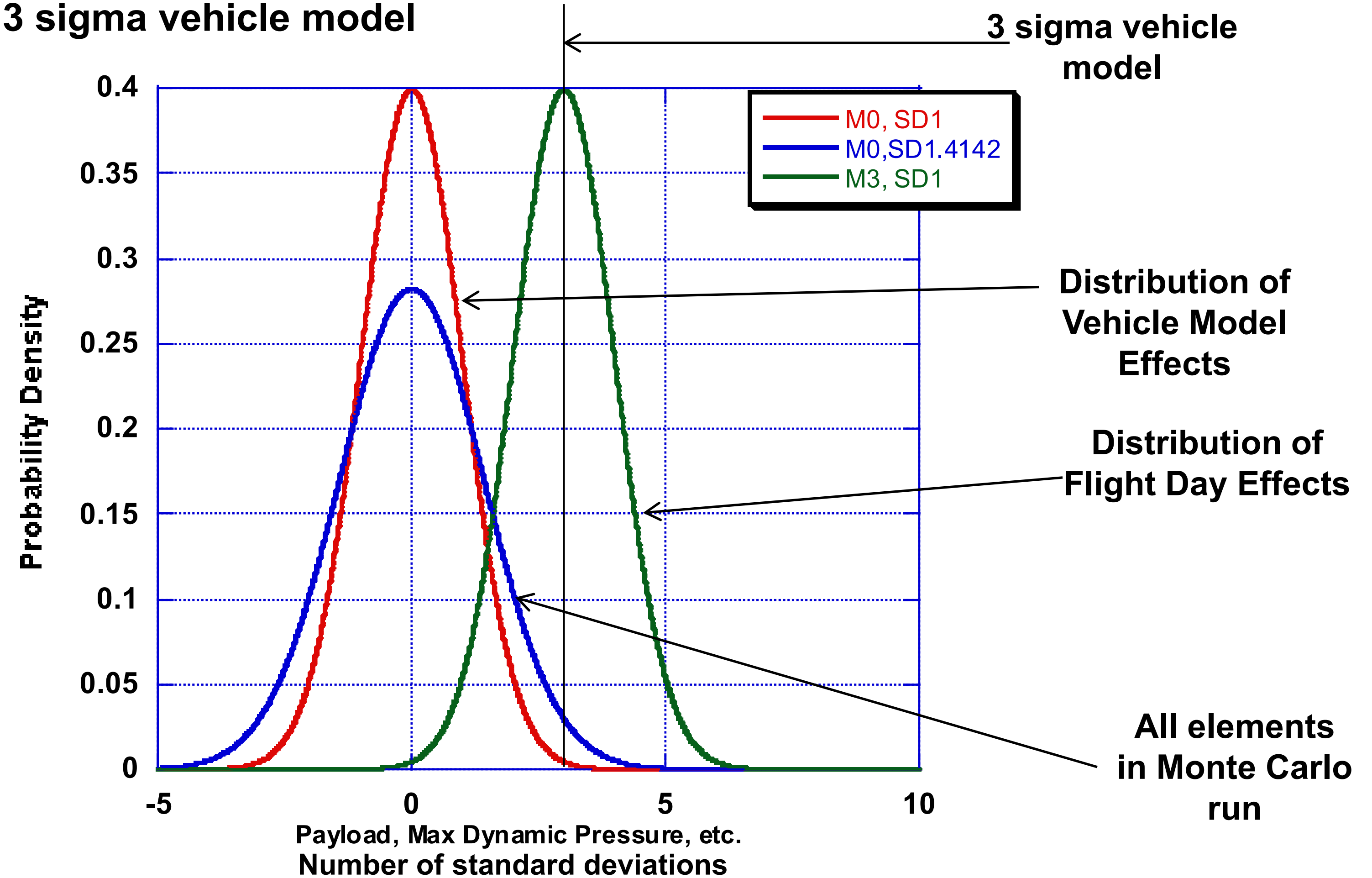
- Four types of variations
  - 1. Uncertainty during the design phases
    - How will the engines perform?
    - What are the aerodynamic parameters?
    - Uncertainty normally decreases with design maturity
  - 2. Know about for the particular assembled vehicle
    - Each manufactured element differs in weight and performance (thrust, Isp, solid burn rate), get estimates of these (for example, if the engine is tested prior to flight)
    - Axial force coefficient will be known better by the time the vehicle flies (wind tunnel testing)
  - Need to be able to meet requirements and design limits with these assembled vehicles, even early in the design phase
  - Design the trajectory in simulation with these as “knowns”

Varying Vehicle Parameters

- Four types of variations
  - 3. Know about on launch day
    - Know for day of launch trajectory design and go/no go decision
      - Temperature estimate (solid rocket motor)
      - Propellant loading estimate
      - Wind estimate
  - 4. Don't know on launch day
    - Remaining uncertainties in all parameters (engine performance, aerodynamics, propellant loading, and more detailed parameters such as actuator modeling)
- If we throw all these into the same Monte Carlo simulation, we won't get the right answer for anything (mission success, flight performance reserve, maximum dynamic pressure, thermal indicators, maximum acceleration, ….)

Vehicle Models vs Flight Day Unknowns

- Why it's critical to separate the unknowns
- Assume both pre-flight known items and flight day unknowns have the same standard deviation and are characterized by Normal distributions
- This example assumes you need to meet mission success requirements when launching a 3 sigma vehicle model



Some Vehicle Parameters of Interest

- Heavy/slow vehicle
  - Lowest payload capability (or lowest remaining propellant)
  - Longest time to clear the tower, most liftoff drift
  - Most lofting, so may be worst abort conditions
- Light/fast vehicle
  - Highest acceleration
  - Highest dynamic pressure
  - Highest bending loads
  - Max temperatures
- Hybrid
  - For some parameters, a different model might be appropriate
    - For example, a booster stage with high performance and a high mass upper stage and payload might give driving compression loads

Steps for the Vehicle Models

- **1. Determine the partial derivatives of the desired varying outputs (*z*) with respect to each of the parameters that can vary (x_i) but will be known for a given assembled vehicle (for example, the partial of dynamic pressure with respect to burn rate).**
  - Use the standard deviation for each parameter, from the uncertainty portion that will be known when the vehicle is assembled
  - Generate the standard deviation for the composite impact of all the parameters

$$\sigma_z = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial z}{\partial x_i} \sigma_i \right)^2}$$

Some parameters will have big impacts, others not so much

- **2. Determine the target (*z*) values of each parameter (e.g. dynamic pressure, heat rate, acceleration, liftoff thrust/weight ratio, low fuel, etc.)**
  - Assuming the composite behaves as a Gaussian (Normal) distribution, we have

$$z = n \, \sigma_z$$

where *n* is the desired sigma level, e.g. $n = 3$ corresponds to 99.865%. We typically use a 90% vehicle model so that we don't have too much conservatism (3 sigma vehicle with 3 sigma flight day variations).

- **3. Determine the vehicle model that uses the total set of parameters and maximizes the statistical chance of ending up with that vehicle.**
  - Constrained numerical optimization
    - Target *z* values are the constraints (could be multiple targets for a vehicle model, e.g. all light/fast target params)
    - x_i values are the variables to be chosen
    - Maximize the joint probability of the x_i from the x_i distributions (can be Normal, Uniform, etc.)
    - The more influential parameters will vary more

Monte Carlo Procedure

- A wind estimate, a propellant temperature estimate (for solid motors), and liquid propellant loading estimates are known when the launch go/no go decision is made.
- Procedure
  - Define the vehicle model to match the *z* targets
  - Run a Monte Carlo simulation to generate the desired percentage of natural environments case
    - e.g. suppose we want to launch in a 95% bad temperature/wind combination
  - Run Monte Carlo simulation for the remaining flight day uncertainties
    - Includes remaining uncertainty in all the parameters used to define the vehicle model
    - Includes remaining environmental uncertainties (e.g. winds)
    - Do this for each stressing vehicle model, launch month, and mission scenario
    - Show requirements are met for these cases
    - Determine parameters such as Flight Performance Reserve required for mission success, as well as design-driving values.
      - Typically hundreds of values of interest

Flight Control Parameters

- **Aleatory uncertainty: it really varies from flight to flight. Example: Winds, solid burn rate**
- **Epistemic uncertainty: there is really only (approximately) a single value, but we don't know it.**
- **The vehicle parameters we have already separated are mathematically epistemic**
  - We don't know them now in general, but will estimate the single value prior to flight. So we need to cover for the variations in the single value. (Mass is a simple example)
- **What about other epistemic values?**
  - e.g. aerodynamic moment coefficients, slosh mode frequency, first vibration mode.
- **These remaining epistemic parameters are primarily impacts to flight control**
  - Flight control effects are highly nonlinear, so it would not be appropriate to just examine the bounds, or specific values (e.g. build a 90% vehicle model), of these parameters and assume all combinations would be covered
- **Not feasible to run a Monte Carlo simulation with varying flight control parameters for each sample of flight day winds and vehicle models**
- **A reasonable approach seems to be:**
  - Monte Carlo run with all parameters included to determine the effects of parameter interactions (use the Monte Carlo runs we already have)
  - Sensitivity studies with individual parameters to see where the system is sensitive and to ensure robustness
  - Margin on flight control stability results

Number of Monte Carlo Samples

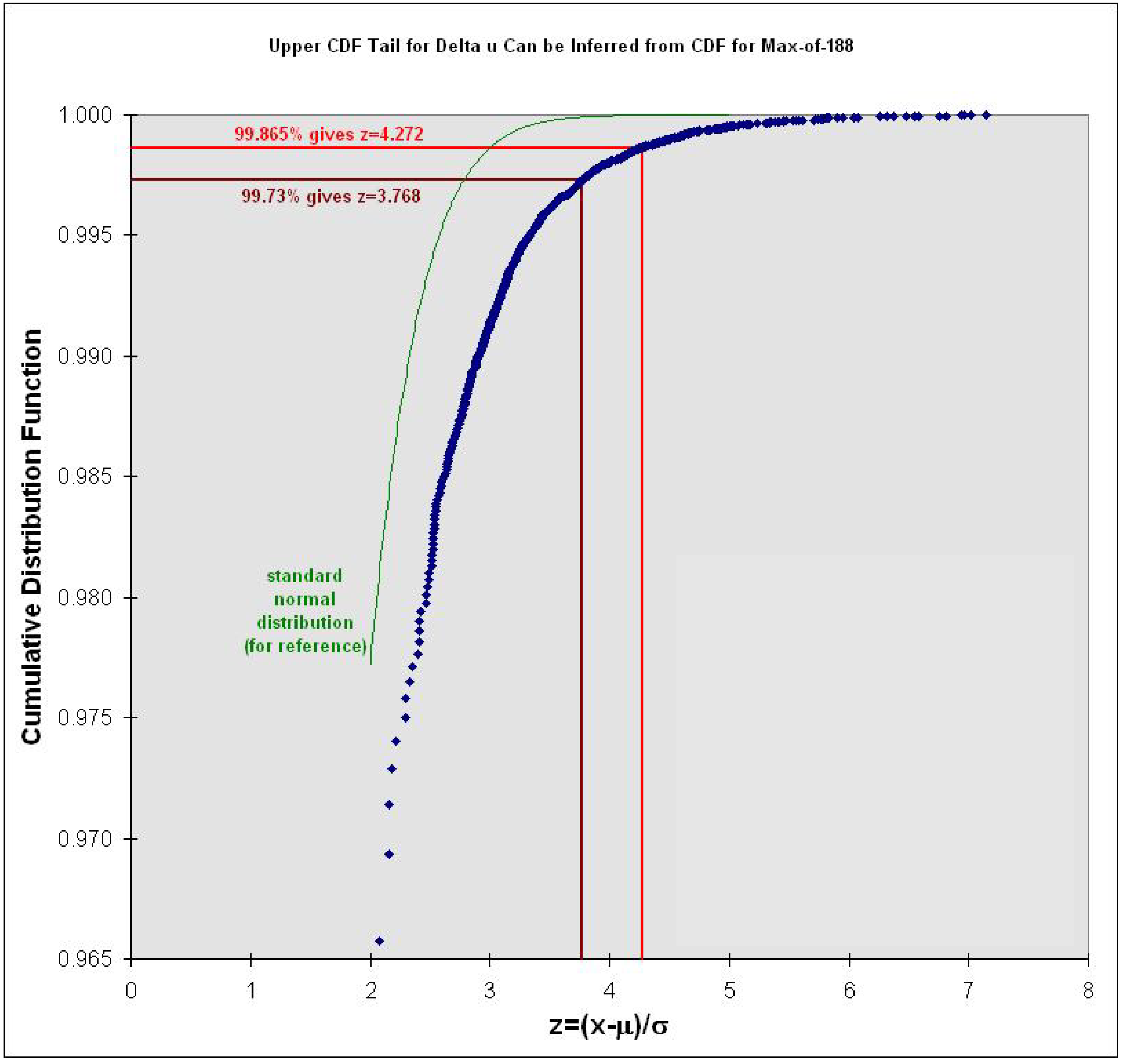
Requirements and Design Values

- Typical requirements statements specify probabilistic success percentages
  - For example, 99.73% with 10% consumer risk (or 90% confidence)
- Or we want to design to a certain percentage worst case
  - For example, 99.865% high maximum acceleration
  - It doesn't make sense to say 100%
- So, how do you capture a certain percentage value from a Monte Carlo simulation?
  - The values we are looking for are among the last few samples in the tails of the data
    - So there will be random variation in these values according to the samples we randomly select
    - If we run a sim with a different random seed, we will get a different answer
- Fitting a distribution to the data is not generally satisfactory because we are interested in the outer few outliers
  - The behavior of these outliers is generally not Gaussian

Example of Why Fitting a Distribution Isn't So Good

Example from data: The tail of these wind profiles is not Gaussian. X-axis is sigma level.

Plotted is the change in easterly wind component from the measured value



Order Statistics

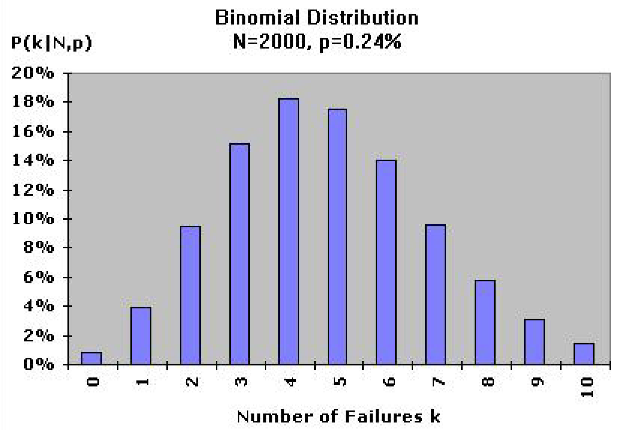
- We develop a method here of addressing requirements success, and determining desired percentage values of parameters, using Order Statistics
  - Order the results from smallest to largest and study the statistics of the outlier cases
- This development is not new to statistics theory
- This is also called "acceptance sampling" in quality engineering

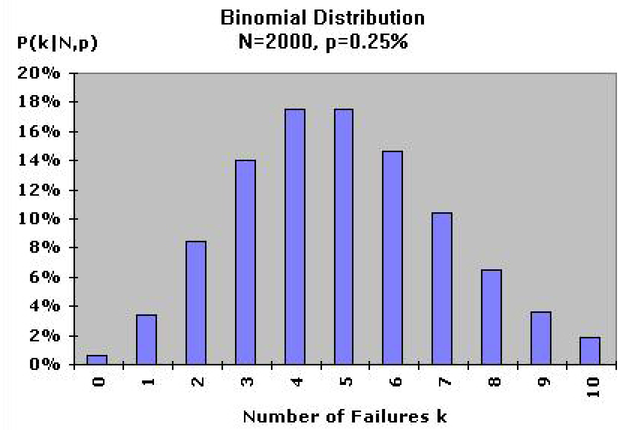
Examining Results of a Monte Carlo Run as Successes and Failures

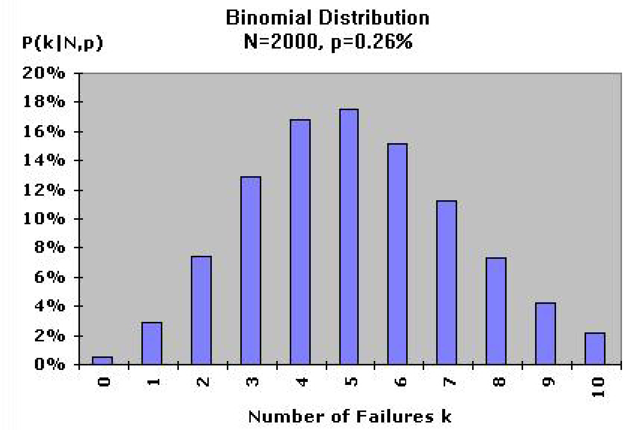
$$P_{\mathrm{BIN}}\left(k \middle| p, N\right) = \binom{N}{k} p^k \left(1-p\right)^{N-k}$$

N **is the number of samples**
p **is the actual failure probability**
k **is the number of failures observed**

If we take 2000 samples, with differing failure probability, how many failures will we see?

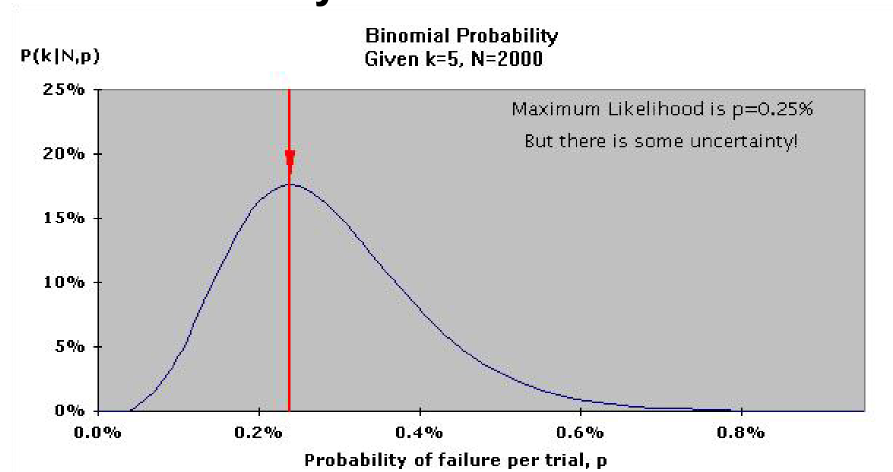




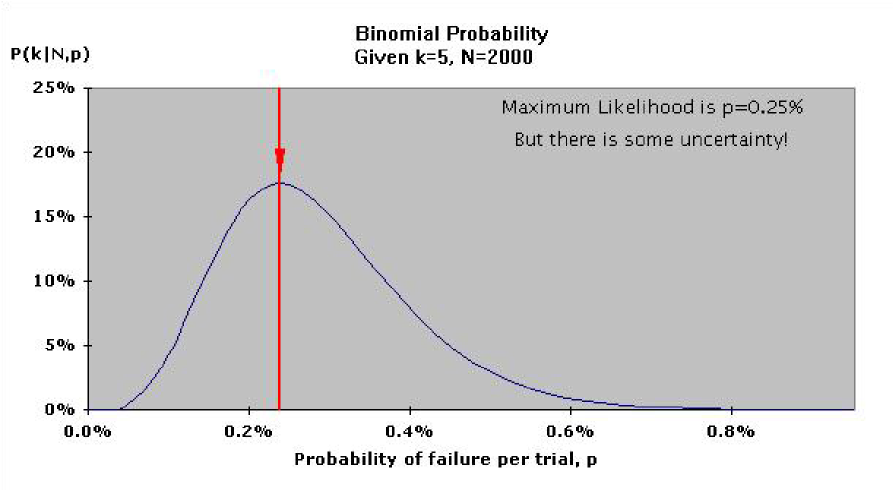


It varies, and the chance of seeing exactly 5 failures is nearly the same in these cases

- **Given 5 failures in 2000 samples (=0.0025*2000), what is the actual probability of failure? It varies.**
- **So, given a required success probability and a confidence level, how do we determine how many failures are allowed?**



Risk for Maximum Likelihood and 10% Risk



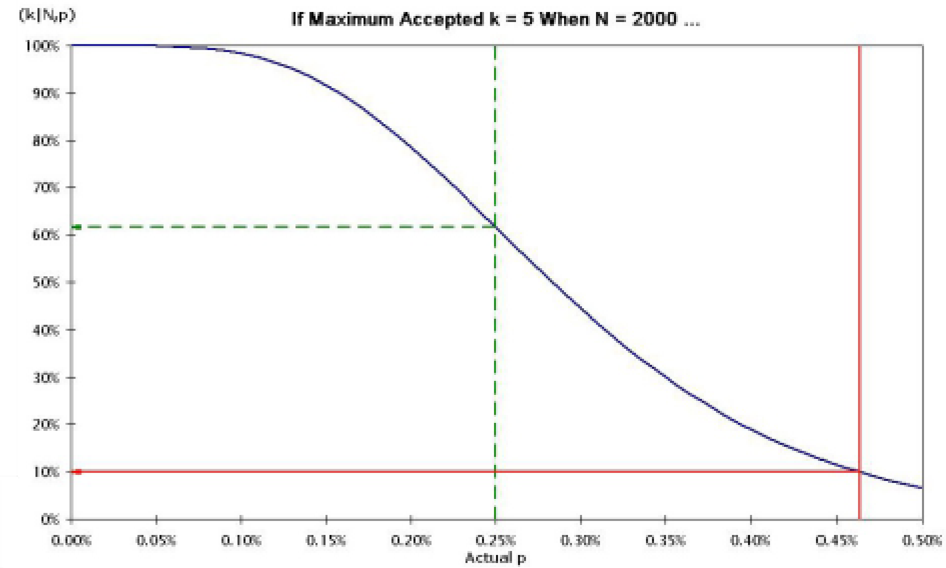


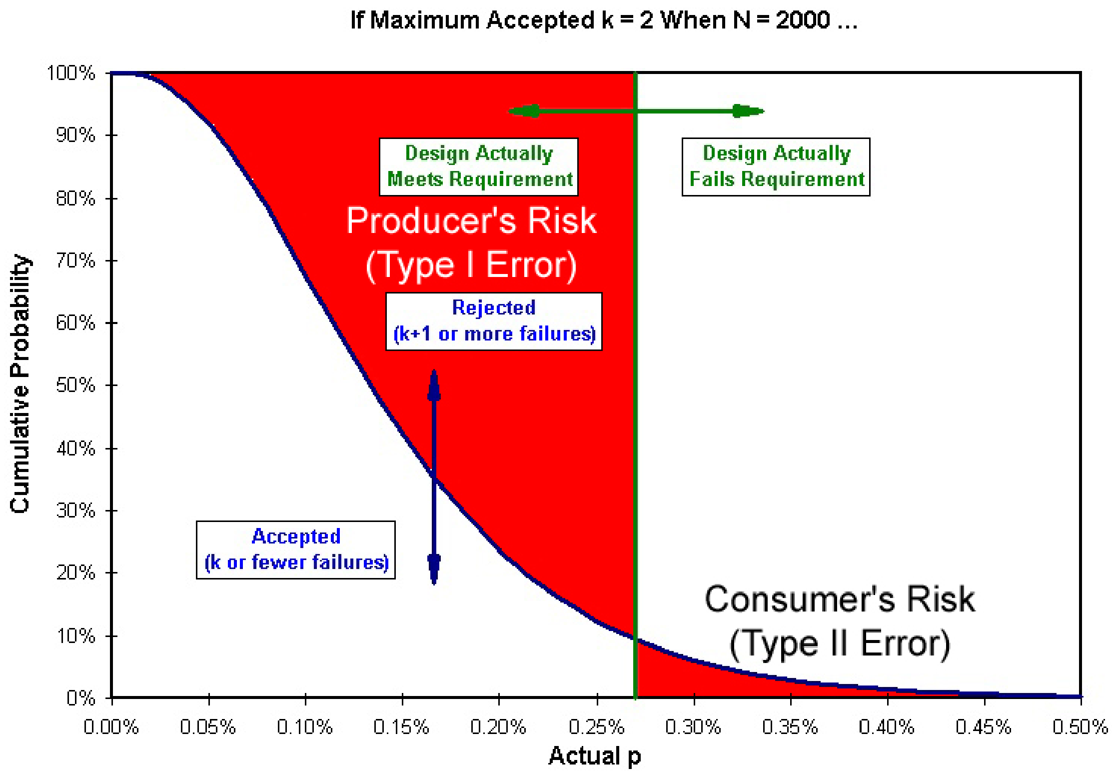
Figure 4. Cumulative probability *F(k|p,N)* as a function of actual failure rate.

- **0.25% is not the 50% risk/confidence level, as may be thought**
- **About a 61% chance that if our requirement is 0.25% failures, the actual failure rate is higher**
- **At about 0.47% failures the risk drops below 10%**

Consumer Risk and Producer Risk

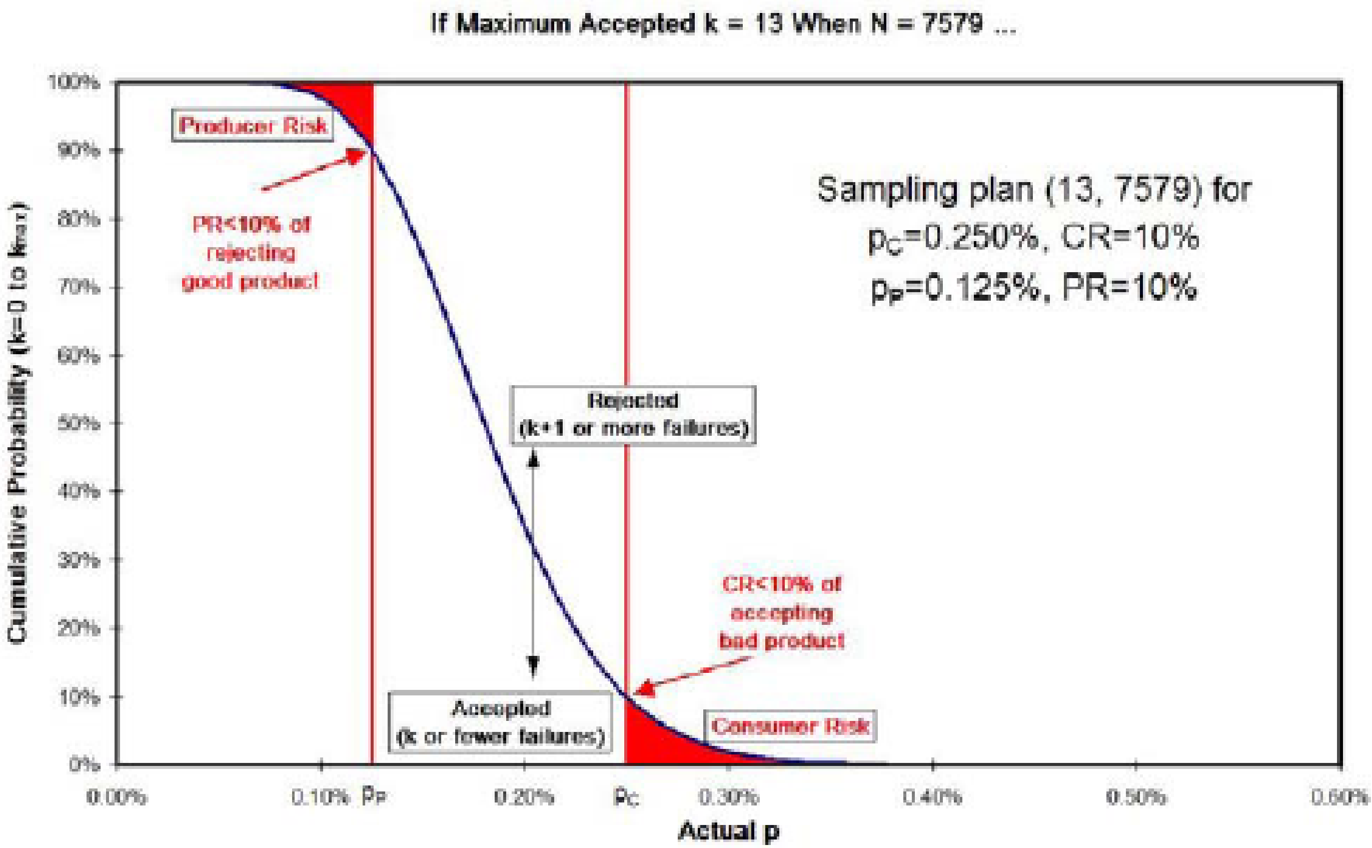
- Required success: 99.73%
- Graphed: probability of seeing 2 or fewer failures if the actual failure probability is as specified on the x-axis
- At 0.27%, there is a 10% chance we would get 2 or fewer failures if we run 2000 samples (and accept a bad product if 99.73% success is the requirement). This is consumer risk.

"Operating Characteristic" curve



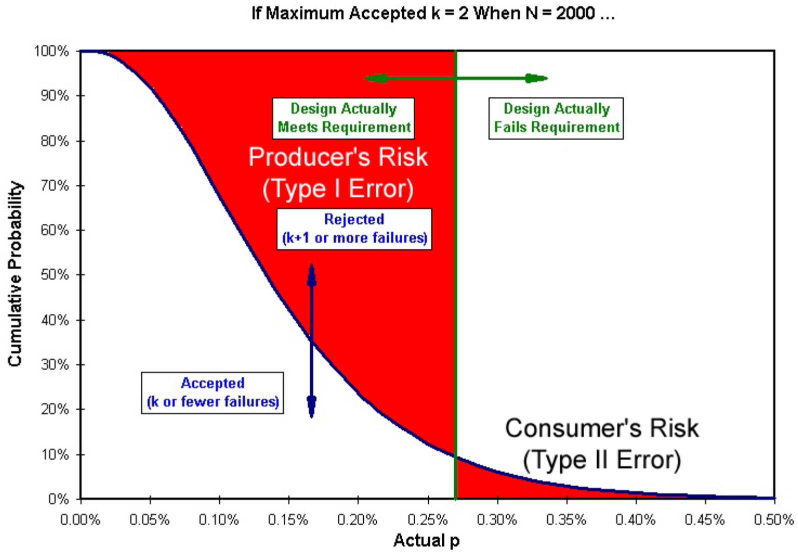
- There is a 90% chance we reject a batch that actually meets the requirement by seeing more than 2 failures (with 99.73% actual success). This is producer risk (rejecting a good product).
- So we design for a lower percentage failure rate in order to reduce the producer risk.

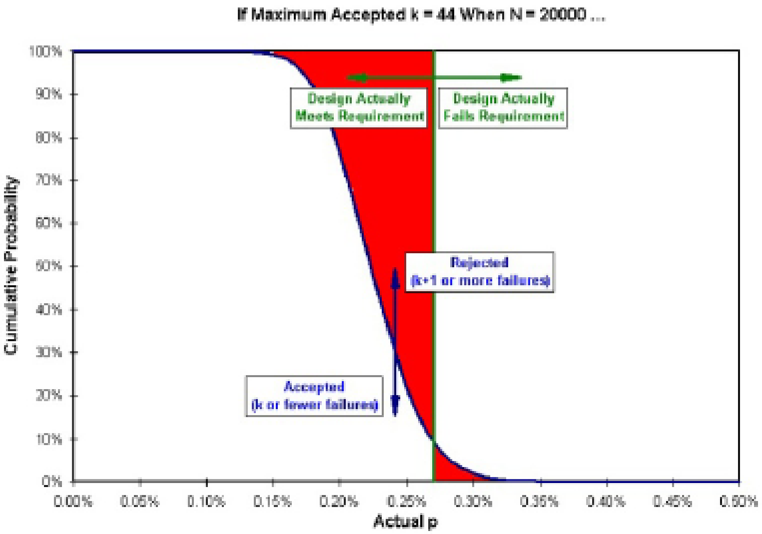
Consumer Risk and Producer Risk

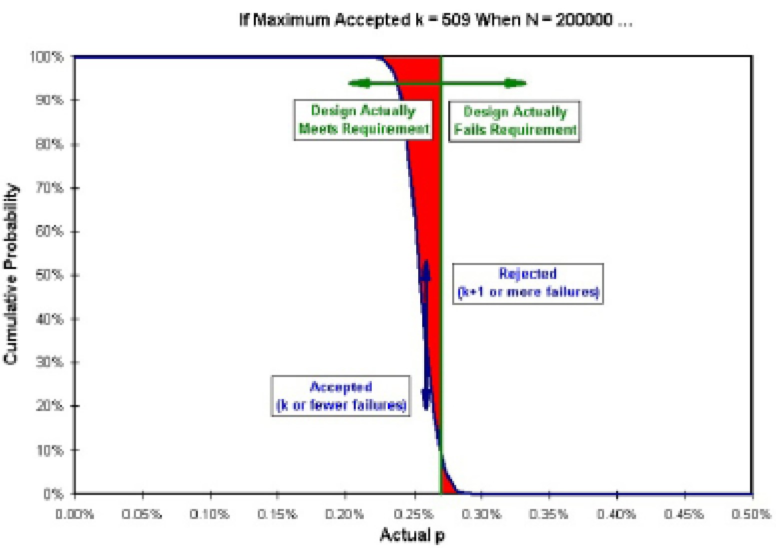


- Design for a lower failure probability to reduce producer risk
- Accept a higher failure probability to reduce consumer risk

Value of More Samples







- **Increasing the number of samples tightens the operating characteristic**
- **Less conservatism needed to meet the 10% consumer risk number**
- **If we ignore producer risk, we can use a smaller number of Monte Carlo samples but need a more conservative design**
- **2/2000 = 0.001, 44/20000 = 0.0022, 509/200000 = 0.002545**

Sample of Number of MC Samples Required

Considering only consumer risk

Failure Fraction Desired	Consumer Risk (%)	Allowable Number of Failures	Minimum Number of Runs Necessary	Allowable Failure Fraction
0.02	10	0	114	0
0.02	10	1	194	0.005155
0.02	10	2	265	0.007547
0.0027	10	0	852	0
0.0027	10	1	1440	0.000694
0.0027	10	2	1970	0.001015
0.00135	10	0	1705	0
0.00135	10	1	2880	0.000347
0.00135	10	2	3941	0.000507
0.00135	10	10	11410	0.000876
0.00135	10	20	20030	0.000999
0.00135	50	0	514	0
0.00135	50	1	1243	0.000805
0.00135	50	2	1981	0.001010
0.00135	50	10	7903	0.001265
0.00135	50	20	15310	0.001306

Measure of conservatism needed

Still much less than 0.02

Getting closer to 0.00135

- **Need more samples if we want to allow for a higher failure fraction (smaller producer risk)**
- **Need more samples for a higher percentage of success**
- **Need more samples for lower consumer risk**
- **If the number of samples possible is limited by expense, we have to accept proving a lower level of success, or higher consumer risk, or both (with higher producer risk)**
- **Number of samples does not depend on the number of input parameters**

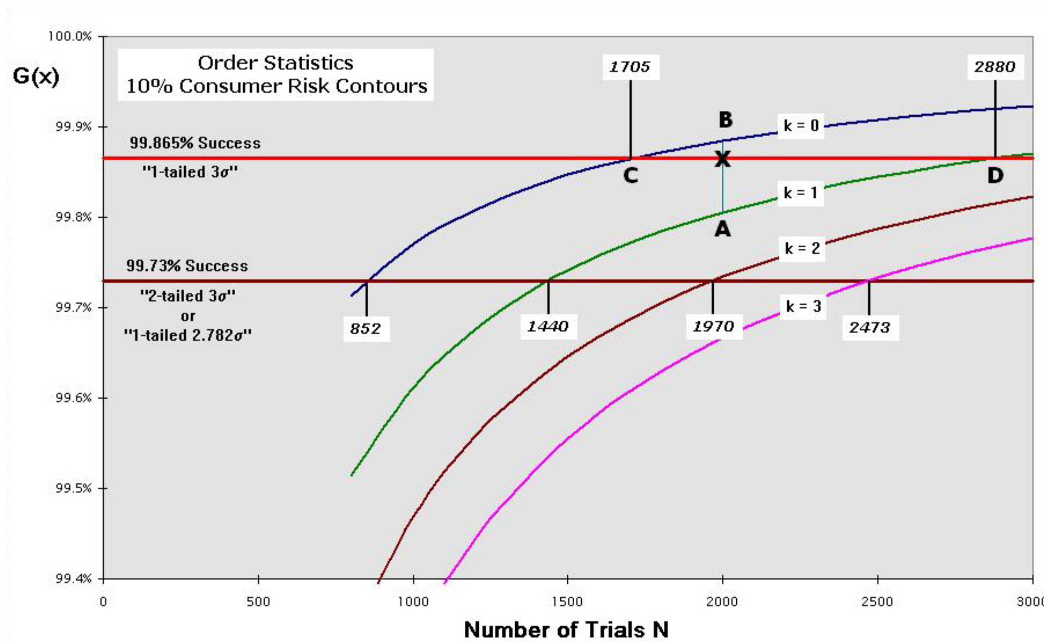
Order Statistics and Design Parameters

- With “Order Statistics”, we simply order our 2000 results from lowest to highest
- The failure cases are those that exceed the limiting value (orbit insertion error, propellant remaining, roll attitude error, stage separation clearance, …..)
- How do we determine the 99.865% with 10% consumer risk maximum dynamic pressure, acceleration, …?
  - Interpolate between the outliers assuming similar triangles

- If 1705 samples, the 99.865%/10% value is equal to the value of point 1705
- Along segment B-A, the 99.865%/10% value of accel is found by interpolating between point 2000 and point 1999

$acc_{1999} + (acc_{2000} - acc_{1999}) *$

$(2880 - 2000) / (2880 - 1705)$



Why this Method is so Useful

- These results are based on binomial successes and failures and are not tied to the underlying distribution
- These results don't depend on how many parameters are varying in the input to the Monte Carlo simulation
- So they hold for all distributions of the input/output data and for any number of inputs

- Be careful however: If you run a lot of samples and think you have the variations well pinned down, they are only as good as the accuracy of the input values and uncertainties.

Summary

- This presentation showed methods for:
  - Separating and analyzing uncertainties that are
    - Known when a vehicle is assembled
    - Known on flight day
    - Unknown on flight day
  - Determining how many Monte Carlo samples are needed in order to verify probabilistic requirements
  - Addressing the level of conservatism versus the number of Monte Carlo samples
  - Determining the value of various design parameters