STI 10-028

An Extreme-Value Approach to Anomaly Vulnerability Identification

Chris Everett[a*], Gaspare Maggio[a], and Frank Groen[b]
[a]Technology Risk Management Operations, ISL, New York, NY
[b]Office of Safety & Mission Assurance, NASA, Washington, D.C.

Abstract: The objective of this paper is to present a method for importance analysis in parametric probabilistic modeling where the result of interest is the identification of potential engineering vulnerabilities associated with postulated anomalies in system behavior. In the context of Accident Precursor Analysis (APA), under which this method has been developed, these vulnerabilities, designated as *anomaly vulnerabilities*, are conditions that produce high risk in the presence of anomalous system behavior. The method defines a parameter-specific Parameter Vulnerability Importance measure (PVI), which identifies anomaly risk-model parameter values that indicate the potential presence of anomaly vulnerabilities, and allows them to be prioritized for further investigation. This entails analyzing each uncertain risk-model parameter over its credible range of values to determine where it produces the maximum risk. A parameter that produces high system risk for a particular range of values suggests that the system is vulnerable to the modeled anomalous conditions, if indeed the true parameter value lies in that range. Thus, PVI analysis provides a means of identifying and prioritizing anomaly-related engineering issues that at the very least warrant improved understanding to reduce uncertainty, such that true vulnerabilities may be identified and proper corrective actions taken.

Keywords: Anomalous Conditions, Accident Precursor Analysis, Anomaly Risk Significance, Importance Measures

1. INTRODUCTION

The objective of this paper is to present a method for importance analysis in parametric probabilistic modeling where the result of interest is the identification of potential engineering vulnerabilities associated with postulated anomalies in system behavior. In the context of Accident Precursor Analysis (APA) [1], under which this method has been developed, these vulnerabilities, designated as *anomaly vulnerabilities*, are conditions that produce high risk in the presence of anomalous system behavior. The method defines a parameter-specific Parameter Vulnerability Importance measure (PVI), which identifies anomaly risk model parameter values that indicate the potential presence of anomaly vulnerabilities, and allows them to be prioritized for further investigation. This entails analyzing each uncertain risk-model parameter over its credible range of values to determine where it produces the maximum risk. A parameter that produces high system risk for a particular range of values suggests that the system is vulnerable to the modeled anomalous conditions, if indeed the true parameter value lies in that range. Thus, PVI analysis provides a means of identifying and prioritizing anomaly-related engineering issues that at the very least warrant improved understanding to reduce uncertainty such that a true vulnerability may be identified and proper corrective action taken.

2. ANOMALOUS CONDITION RISK

APA serves as the bridge between existing risk modeling activities, which are often based on historical or generic failure statistics, and system anomalies, which provide crucial information about the failure mechanisms that are actually operative in the system. The APA technical approach entails the development of a parametric probabilistic anomaly risk model for risk significant anomalies that can be exercised in a number of ways to generate the risk results of interest. A primary result is the Anomalous Condition Risk (ACR), which is the conditional risk that is directly attributable to a failure mechanism occurring outside nominal bounds, thereby creating an anomalous condition. ACR is

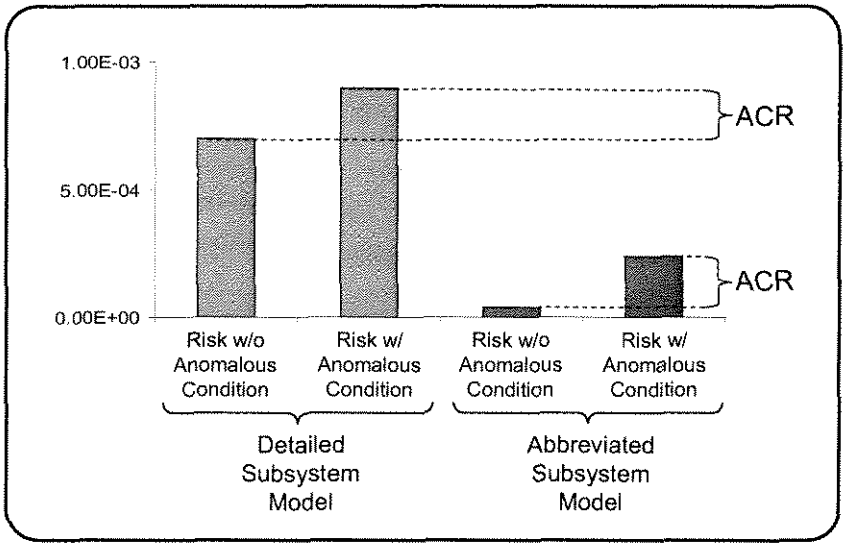
*Email address for contact author: ceverett@islinc.com

conditional, i.e. it presumes the occurrence of an anomaly, but the particular characteristics of the anomaly are indeterminate. This means that ACR is calculated by allowing the anomaly failure mechanism to vary (in magnitude, location, etc.) in accordance with the failure mechanism's parent distribution, and the expected risk over all resulting anomalies is calculated as:

$$\text{ACR} = (\text{Risk} \mid \text{Anomalous Condition}) - (\text{Risk} \mid \sim\text{Anomalous Condition}) \quad (1)$$

The second term in the Equation 1 is needed to remove any stochastic risk that is represented in the model but not attributable to the anomaly, such as that due to random failures of modelled safety systems.[†] This is done not only to isolate the risk that is directly attributable to the anomaly event, but also to remove the effects of variability in model scope. For example, a detailed model that includes all the components of a subsystem will typically show more risk than a subsystem model that is restricted to the components directly involved in the anomaly. The situation is shown in Figure 1. The ACR is used in the APA technical approach to prioritize anomalies in terms of their risk significance, and offers a measure that can be used to establish criteria for precursor designation.

Figure 1: Anomalous Condition Risk (ACR) for the Same Anomaly but Derived from Models of Varying Scale



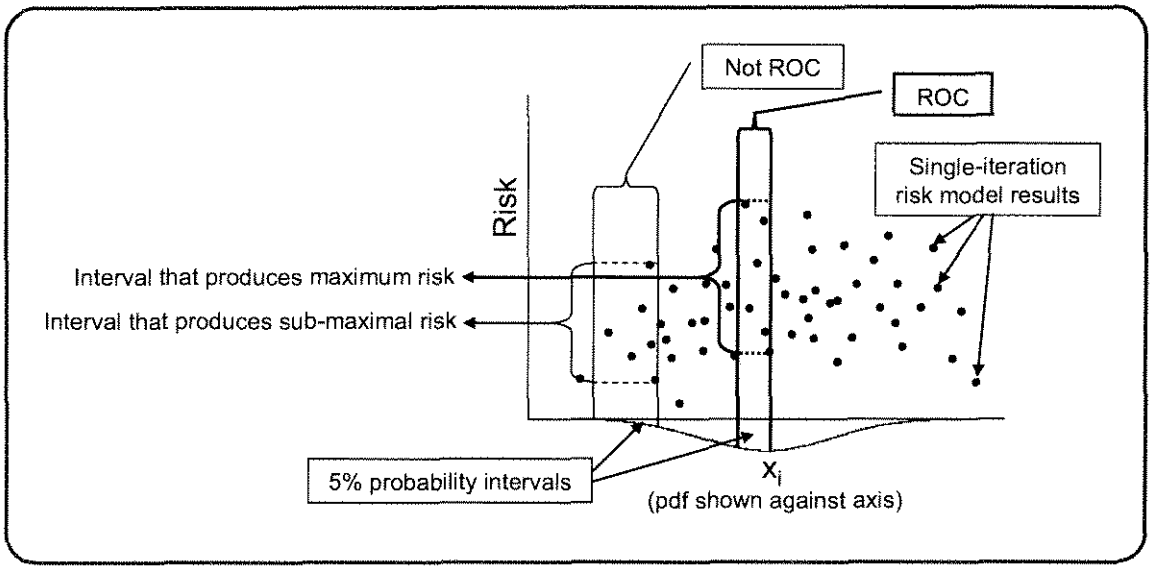
3. PARAMETER EXTREME RISK

ACR addresses the expected risk to the system conditional on anomaly occurrence, but it does not provide direct insight into potential anomaly-specific vulnerabilities which should be investigated further. To accomplish this, a parameter-specific risk metric, Parameter Extreme Risk (PER), is defined for each uncertain parameter x_i in the anomaly risk model, involving the recalculation of ACR under the restriction that the x_i lies in that region of its probability density function (pdf) that is of greatest concern. The precise definition of this region is somewhat arbitrary, but for the purpose of calculating PER the *region of concern* (ROC) for parameter x_i is defined as the 5% probability interval over x_i that produces the largest risk, given an anomalous condition. For example, if parameter x_i represents a failure threshold whose location is uncertain and therefore described by a pdf over x_i, the ROC would be the left-hand tail of the x_i pdf, since that is where the failure threshold is at its lowest value, producing the greatest vulnerability to imposed stresses. Thus, in this case it can be determined by inspection that the ROC is the region to the left of the 5th percentile value of x_i. In the context of APA, which involves first-order probabilistic parametric modeling of anomalous-condition-induced failure scenarios, it is anticipated that the ROC will usually be determinable by inspection based on physical arguments. When this is not the case, the ROC can be determined by finding the 5%

[†] A mathematical justification for taking the difference between these probabilities is given in Appendix A.

probability interval over x_i that produces the highest risk, given an anomalous condition, using the model results generated for the purpose of calculating ACR[‡]. This is shown schematically in Figure 2. The individual Monte Carlo trials that fall within the anomalous range of the anomaly failure mechanism are plotted in terms of x_i and risk. For every 5% probability interval over x_i, the trials that lie within are averaged to produce a risk for that interval. The interval for which the risk is at a maximum is the ROC.

Figure 2: Determination of the Region of Concern Risk Model Monte Carlo Results



Once the ROC has been determined, PER for x_i is calculated as:

$$PER_i = (\text{Risk} \mid \text{Anomalous Condition} \wedge x_i \in \text{ROC}) - (\text{Risk} \mid \sim\text{Anomalous Condition}) \qquad (2)$$

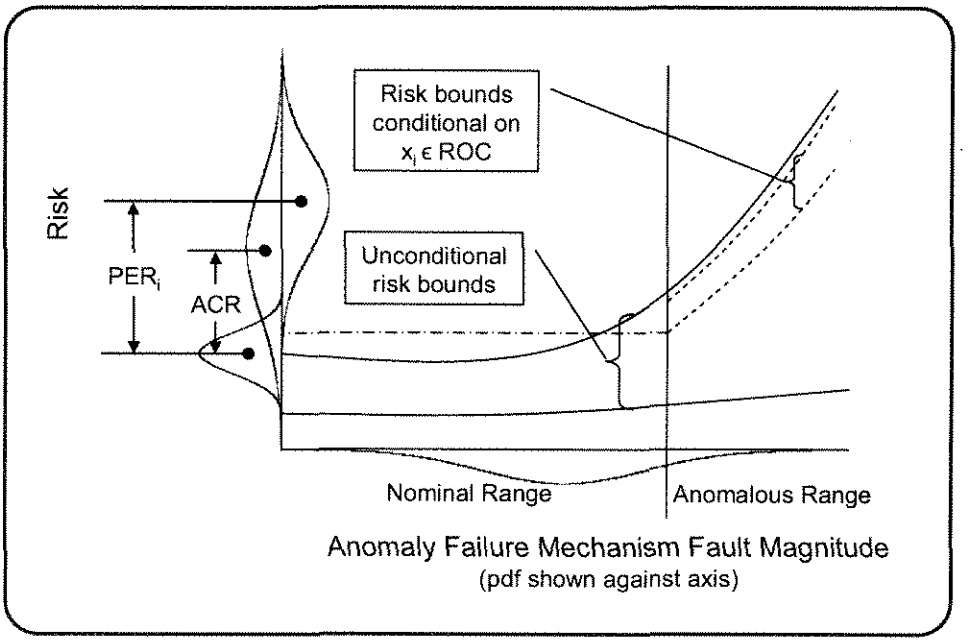
The situation is shown schematically in Figure 3. The magnitude of the anomaly failure mechanism is shown on the x-axis as a pdf, the right hand tail of which lies in the anomalous range. For a given value of the anomaly failure mechanism's fault magnitude, the risk model produces a risk pdf representing the range of possible risk over the distributions of uncertain parameters in the model. The solid lines represent the risk range when the parameters are unconstrained, and the dashed lines represent the risk range when parameter x_i is constrained to the ROC. The three pdfs on the risk axis represent the conditional risk distributions that factor into the calculation of ACR and PER_i. The lower pdf on the left side of the y-axis is the risk conditional on the anomaly failure mechanism being within its nominal range (i.e. nominal risk). The upper pdf on the left side of the y-axis is the risk conditional on the anomaly failure mechanism being within the anomalous range. The difference between the expected values of these two pdfs is the ACR, which measures the risk that is attributable to the anomaly failure mechanism occurring within the anomalous range. The pdf on the right side of the y-axis is the risk conditional on the anomaly failure mechanism being within the anomalous range *and* on parameter x_i lying within the ROC. Its minimum value is demarcated by the hashed horizontal line. The difference between its expected value and that of the nominal risk pdf is the PER_i.

Like ACR, PER is calculated as the difference between two risks. However, in this case, subtraction of the second term on the right hand side of Equation 2 does not leave just the risk that is attributable to the anomalous condition. It also leaves the risk from non-anomalous-condition sources that is attributable to x_i being in the ROC. This is because the second term is the risk, given non-occurrence of the anomalous condition, over the entire x_i pdf, not just over the ROC. Therefore, if PER is high, it

[‡] Since, by definition, PER calculation will use (roughly) 5% of the total number of risk model runs, the number of Monte Carlo iterations of the risk model required for statistical stability should be determined accordingly.

might be because the system is particularly vulnerable to the anomalous condition in the ROC, or because the system is inherently vulnerable in the ROC even in the absence of the anomalous condition.

Figure 3: Schematic Representation of ACR and PER_i.



To distinguish these two sources of ROC-related risk, PER_i can be decomposed into two parts: PER_i^{AO}, which addresses that part of PER_i that is attributable to the anomalous condition, and PER_i^{NA}, which addresses that part of PER_i that is attributable to other sources.

$$PER_i^{AO} = (\text{Risk} \mid \text{Anomalous Condition} \wedge x_i \in ROC) - (\text{Risk} \mid \sim\text{Anomalous Condition} \wedge x_i \in ROC) \quad (3)$$

$$PER_i^{NA} = (\text{Risk} \mid \sim\text{Anomalous Condition} \wedge x_i \in ROC) - (\text{Risk} \mid \sim\text{Anomalous Condition}) \quad (4)$$

PER_i^{AO} and PER_i^{NA} sum to PER_i, and together provide a means of assessing the significance of the anomalous condition in contributing to an increase in risk, over and above the nominal risk, for the situation where a parameter is within its ROC[§].

PER has dimensions of risk, and in fact is an explicit representation of risk, conditional on the parameter-specific constraints. Its calculation can be thought of as a form of risk-based "what if" analysis – for each parameter it answers questions of the type: "*What* is the risk, *if* the parameter is in a range that produces a vulnerable system?" A high value of PER indicates a potential vulnerability, associated with parameter x_i, due to either the anomalous condition or to inherent system risk, which might merit additional investigation to determine whether or not the actual value of the parameter lies in the vulnerable range. Thus, PER raises specific engineering concerns within a risk-based assessment framework that prioritizes them in terms of their overall risk-significance.

[§] If the parameter is the one whose anomalous occurrence is under investigation, it is possible, and indeed not improbable, for the ROC to be found wholly within the anomalous region of the parameter's pdf. In this case, (Risk | ~Anomalous Condition $\wedge$ $x_i \in ROC$) is indeterminate because (~Anomalous Condition $\wedge$ $x_i \in ROC$) is the null set. To handle this special case, the following definitions are used:

$$PER_i^{AO} = PER_i; \text{ and}$$
$$PER_i^{NA} = 0.$$

It is worth noting that although PER indicates the possible magnitude of the anomalous condition risk, if x_i is indeed in the ROC, it does not indicate the range of x_i values for which the risk is at or near its PER value. It is possible that this range is larger than the ROC, or even that there is more than one region of the pdf where risk is high. Thus, the method is exhaustive in terms of identifying all parameters associated with potential anomalous condition vulnerabilities, but it is not exhaustive in terms of identifying all values of a given parameter for which a potential vulnerability may exist.

4. IMPORTANCE MEASURES

Both ACR and PER represent explicit conditional risk contributions to an overall system risk, and as such they are most meaningful in the context of that risk. If the risk at the overall system level is R_o, then the ratio of ACR can be taken with respect to R_o to define the following importance measure:

$$\text{Anomalous Condition Risk Importance (ACRI):} \quad ACRI = \frac{ACR}{R_o} \tag{5}$$

Once an anomaly has been designated as important, based on ACRI, the following parameter-specific importance measure is defined in order to indicate the potential presence of a parameter-related vulnerability:

$$\text{Parameter Vulnerability Importance } (PVI_i)\text{:} \quad PVI_i = \frac{PER_i}{ACR} \tag{6}$$

PVI can be further decomposed, using PER^{AO} and PER^{NA}, into PVI^{AO} and PVI^{NA}, where PVI^{AO} measures the parameter vulnerability due to the anomalous condition, and PVI^{NA} measures the vulnerability due to the intrinsic behavior of the system.

In the case of ACRI, normalization with respect to the overall system risk allows the significance of anomalous conditions to be assessed relative to other risks in the system, which supports prioritization of risk management attention among competing issues. It also provides a system-independent means for designating an anomaly as a precursor, in cases where precursor criteria are established. For example, an ACRI value of 1% or greater could be considered a reasonable basis for precursor designation. In the case of PVI, normalization with respect to ACR supports prioritization of potential vulnerabilities across different parameters, for anomalous conditions whose ACRI values warrant a parameter-specific level of investigation. By definition, PVI must be greater than or equal to one. A PVI value at or near one indicates that changes in the corresponding parameter do not produce significant changes in anomalous condition risk. Larger values of PVI indicate the presence of parameter ranges that may be risk drivers. These risk drivers may be related to anomalous conditions or they may be intrinsic to the system, regardless of the anomalous condition. In general, the absolute value of a parameter's PVI is less important than its value relative to other parameters, since its purpose it to help prioritize potential engineering issues for investigation, rather than to give an absolute indication of risk.

It is worth noting that both ACRI and PVI are calculated using risks that are conditional on the occurrence of the anomalous condition. Therefore, when comparing measures across anomalous conditions, the probabilities of occurrence of the conditions are not taken into account. The measures have been defined in this way because anomaly investigation and precursor analysis are intrinsically pre-emptive, i.e. the intent is ideally to find and eliminate vulnerabilities upon their first manifestation. While this isn't necessarily achievable in all cases, it means that the measures will maintain their effectiveness in a data-lean environment, i.e. before many anomalies have occurred.

When a system risk model exists, the system risk R_o can be taken from the model. In cases where a system risk model does not exist, R_o must be obtained by other means in order to provide a

normalization factor against which risk significance can be established. One possible basis for normalization, in the absence of a calculated risk, is the system risk requirement, which establishes a de facto acceptable risk. Note that assuming an R_o of 0 (meaning the system has absolutely no risk) leads to an ACRI of infinity – this makes sense since a system with no risk should not be experiencing anomalies in the first place. It is imperative that the same benchmark value of R_o be used in calculating ACRI for all anomalous conditions in a particular system, since this is a comparative measure that is intended for use in prioritizing the expenditure of resources to decrease system risk. Therefore, if the benchmark is updated, it should be applied retroactively to all previous ACRI estimates to ensure a common basis.

5. RELATION TO STANDARD RISK AND UNCERTAINTY MEASURES

As part of the development of ACRI and PVI, a number of standard risk and uncertainty measures were reviewed for applicability to anomaly vulnerability identification [2]. Many standard measures are specific to probabilistic risk assessment (PRA) basic events, which presents two distinct difficulties when applying them to model parameters. First, measures that are based on turning events "on" or "off" (i.e. P(event) = 1 or P(event) = 0) cannot be directly applied to parameters, since model parameters do not necessarily have absolute maxima or minima. The Fussel-Vesely (F-V), Risk Reduction Worth (RRW), and Risk Achievement Worth (RAW) are examples of this kind of measure. Secondly measures that are differential cannot be directly applied because there is not necessarily a common basis among parameters with respect to which the differential can be defined. This is not a problem for basic events, which are all probabilities by definition. Examples of this kind of measure include the Birnbaum Measure and the Differential Importance Measure (DIM) under criterion H1[**].

The DIM under criterion H2[††] is mathematically compatible with parametric probabilistic modeling but is not designed to address the specific concerns of anomaly vulnerability identification. It is a local measure, so it does not explicitly explore parameters at their bounding values, where threshold effects are most likely to represent vulnerabilities. Also, its differentials are scaled in terms of the value of the parameter, which has no intrinsic relevance to precursor analysis and its intimate relationship to uncertainty. Entropy-based approaches were also considered, but were determined to be too indirect and potentially very sensitive to assumptions concerning the functional forms of uncertainty characterization.

Although no reviewed measure was entirely adequate to the task of identifying potential anomaly-related engineering vulnerabilities, PVI can be viewed as an adaptation of the RAW measure. RAW is calculated using the following expression:

$$I_{x_i}^{RAW} = \frac{R \mid Pr(x_i) = 1}{R_o} \tag{7}$$

It represents the increase in expected risk conditional on basic event x_i being at its bounding value of unity. Similarly, PVI is conditional on parameter x_i being within its bounding range of values (in terms of resultant risk).

It is of critical importance for the risk model to identify parameter uncertainty where it exists, using a conservative, evidence-based approach. This is because of the concern that vulnerabilities may exist due to overconfidence in the capability of the system to function under stress, particularly when that stress is outside the bounds of operational and/or test experience. Since this is precisely the realm of APA, it is imperative to minimize the reproduction of any overconfidence that might be present in existing system models, erring on the side of conservatism when system behavior is not supported by evidence. Given this approach, the analysis does not *positively* indicate the presence of a vulnerability;

[**] Criterion H1 assumes a uniform change for all parameters (i.e., $\delta x_i = \delta x_j$).

[††] Criterion H2 assumes a uniform percentage change for all parameters (i.e., $\delta x_i/x_i = \delta x_j/x_j = \omega$).

instead, it identifies areas where the available evidence doesn't *rule out* the vulnerability. The onus is then on the system operator to either show that the potential vulnerability is not involve credible conditions, or to address it if it does. This is in keeping with the safety philosophy that puts the burden of proof on the operator to show that the system is safe, rather than on the safety organization to show that it is unsafe. In any case, the issue of whether or not the vulnerable range of a parameter is credible should not be left to chance.

6. SUMMARY

In summary, this paper presents an approach to anomaly vulnerability identification and defines the ACRI and PVI importance measures, which support the identification of potential engineering vulnerabilities that could lead to system failure under anomalous conditions. ACRI addresses the risk-significance of the anomaly's underlying failure mechanism, and supports prioritization of further anomaly investigation and the establishment of criteria for precursor designation. PVI decomposes into PVI^{AO}, which measures the potential vulnerability of the system to the parameter under the anomalous condition of interest, and PVI^{NA}, which measures the potential vulnerability under non-anomalous conditions.

APPENDIX A: ANALYTICAL BASIS OF ANOMALOUS CONDITION RISK

Let the total probability of failure of a system be P(fail); let the condition where a specific anomalous condition exists be AC; and let the situation where a non-anomalous-condition-related failure-causing condition exists in the system be Q. Note that although Q implies that failure is inevitable, it does not imply that the failure will necessarily be due to Q, since a system with both Q and AC might fail due to one or the other condition.

By the law of total probability:

$$\begin{aligned} P(\text{fail}) = {} & P(\text{fail} \mid AC \wedge Q) \times P(AC \wedge Q) + \\ & P(\text{fail} \mid AC \wedge \sim Q) \times P(AC \wedge \sim Q) + \\ & P(\text{fail} \mid \sim AC \wedge Q) \times P(\sim AC \wedge Q) + \\ & P(\text{fail} \mid \sim AC \wedge \sim Q) \times P(\sim AC \wedge \sim Q). \end{aligned} \tag{8}$$

In principle, the risk attributable to the anomalous condition, conditional on its occurrence, is $P(\text{fail} \mid AC \wedge \sim Q)$. However, whereas it is practical to construct a risk model that allows the results to be conditioned on AC vs. ~AC, it is not practical to construct a risk model that allows results to be conditioned on Q vs. ~Q. Thus, the quantities that are amenable to calculation are:

$$\begin{aligned} P(\text{fail} \mid AC) = {} & P(\text{fail} \mid AC \wedge Q) \times P(Q) + \\ & P(\text{fail} \mid AC \wedge \sim Q) \times P(\sim Q) \end{aligned} \tag{9}$$

and

$$\begin{aligned} P(\text{fail} \mid \sim AC) = {} & P(\text{fail} \mid \sim AC \wedge Q) \times P(Q) + \\ & P(\text{fail} \mid \sim AC \wedge \sim Q) \times P(\sim Q). \end{aligned} \tag{10}$$

Now, since Q implies failure, and $\sim AC \wedge \sim Q$ implies success,

$$P(\text{fail} \mid AC \wedge Q) = P(\text{fail} \mid \sim AC \wedge Q) = 1 \tag{11}$$

and

$$P(\text{fail} \mid \sim AC \wedge \sim Q) = 0. \tag{12}$$

Substituting Equations 11 and 12 into Equations 9 and 10 yields:

$$P(\text{fail} \mid AC) = P(Q) + P(\text{fail} \mid AC \wedge \sim Q) \times P(\sim Q) \tag{13}$$

and

$$P(\text{fail} \mid \sim AC) = P(Q). \tag{14}$$

Taking the difference yields:

$$P(\text{fail} \mid AC) - P(\text{fail} \mid \sim AC) = P(\text{fail} \mid AC \wedge \sim Q) \times P(\sim Q). \tag{15}$$

Since the left hand side of Equation 15 is the definition of anomalous condition risk (ACR), we have:

$$ACR = P(\text{fail} \mid AC \wedge \sim Q) \times P(\sim Q). \tag{16}$$

The method for calculating ACR accepts an error of $P(\sim Q)$ relative to $P(\text{fail} \mid AC \wedge \sim Q)$, which does not make a practical difference as long as $P(\sim Q)$ is close to 1, i.e. as long as $P(Q) << 1$. This makes intuitive sense, because subtracting out the fraction of cases with condition Q also subtracts out those cases with condition $Q \wedge AC$, some of which might fail due to AC. Thus the method undercounts failures due to AC in proportion to the fraction of total cases that have condition Q.

Acknowledgements

The authors would like to acknowledge the support of Dr. Michael Stamatelatos and Dr. Homayoon Dezfuli of NASA's Office of Safety and Mission Assurance (OSMA). Anthony Hall of ISL's Technology Risk Management Operation (TRMO) contributed to the refinement of the approach by applying it to selected Space Shuttle anomalies as part of the overall APA process development effort. Scott Insley, also of ISL's TRMO, is acknowledged for his editorial contributions to this document.

References

[1] F. Groen, M. Stamatelatos, H. Dezfuli, and G. Maggio. "*An Accident Precursor Analysis Process Tailored for NASA Space Systems*," 10th International Probabilistic Safety Assessment and Management Conference, PSAM-10, Seattle, WA. June 2010.

[2] M. Stamatelatos, "*Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners*," National Aeronautics and Space Administration, 2002, Washington, D.C.

[3] C.K. Park and R.A. Bari, "*An Information-Theoretic Approach to Uncertainty Importance*," BNL-NUREG-352, Brookhaven National Laboratory, 1985.