Gear Health Threshold Setting Based On a Probability of False Alarm

Eric Bechhoefer[1], David He[2], and Paula Dempsey[3]

[1] *NRG Systems, Hinesburg, VT, 05461*
erb@nrgsystems.com

[2] *University of Illinois at Chicago, Department of Mechanical and Industrial Engineering, Chicago, IL, 69697, USA*
davidhe@uic.edu

[3] *NASA, Glenn Research Center, Cleveland, OH, 69697, USA*
paula.j.dempsey@nasa.gov

ABSTRACT

There is no established threshold or limit for gear vibration based condition indicators (CI) that indicates when a gear is in need of maintenance. The best we can do is set CI thresholds statistically, based on some small probability of false alarm. Further, to the best of our knowledge, there is no single CI that is sensitive to every failure mode of a gear. This suggests that any condition based maintenance system for gears will have some form of sensor fusion.

Three statistical models were developed to define a gear health indicator (HI) as a function of CI: order statistics (max of *n* CIs), sum of CIs and normalized energy. Since CIs tend to be correlated, a whitening process was developed to ensure the HI threshold is consistent with a defined probability of false alarm. These models were developed for CIs with Gaussian or Rayleigh (skewed) distributions. Finally, these functions, used to generate HIs, were tested on gear test stand data and their performance evaluated as compared to the end state of the gear (e.g. photos of damage). Results show the HIs performed well detecting pitting damage to gears. [*]

1 INTRODUCTION

Vibration based gear fault detection algorithms have been developed to successfully detect damaged on gears (McFadden and Smith 1985). Significant effort has also been expended to validate the efficacy of these algorithms (Zakrajsek 1993, Lewicki *et al*. 2010). These studies have demonstrated the ability of gear CI algorithms to detect damage. However, they have not established standardized threshold values for a given quantified level of damage. Additionally, it has been shown (Wemhoff *et al*. 2007, Lewicki *et al*. 2010) that different algorithms are sensitive to different fault modes (Tooth Crack, Tooth Spacing Error, Tooth surfacing Pitting).

The concept of thresholding was explored by Byington et al. (2003), where for a given, single CI, a Probability Density Function (PDF) for the Rician/Rice statistical distribution was used to set a threshold based on an probability of false alarm (PFA). No single CI has been identified that works with all fault modes. This suggests that any functioning condition monitoring will use *n* number of CIs in the evaluation of gear health. A need exists for a procedure to set a PFA for a function using *n* number of CIs.

All CIs have a probability distribution (PDF). Any operation on the CI to form a health index (HI), is then a function of distributions (Wackerly *et al*. 1996). Functions such as:

- The maximum of *n* CI (the order statistics)
- The sum of *n* CIs, or
- The norm of *n* CIs (energy)

are valid if and only if the distribution (e.g. CIs) are independent and identical (Wackerly *et al*. 1996). For Gaussian distribution, subtracting the mean and dividing by the standard deviation will give identical *Z* distributions. The issue of independence is much more difficult.

Two CIs are independent if the probability (P) of CI_1 and CI_2 are equal to:

$$P\left(CI_1 \cap CI_2\right) = P\left(CI_1\right)P\left(CI_2\right) \qquad (1)$$

Equivalently, CI_1 and CI_2 are independent random variables if the covariance of CI_1, CI_2 is 0. This is, in general, not the case, where the correlation coefficient is defined as the covariance divided by the standard deviation:

$$\rho = {Cov\left(CI_1, CI_2\right)} \big/ {\sigma_1 \sigma_2} \qquad (2)$$

The range of correlation coefficients used in this study for pairs of gear CIs are listed in Table 1.

ρ_{ij}	**CI 1**	**CI 2**	**CI 3**	**CI 4**	**CI 5**	**CI 6**
CI 1	1	0.84	0.79	0.66	-0.47	0.74
CI 2		1	0.46	0.27	-0.59	0.36
CI 3			1	0.96	-0.03	0.97
CI 4				1	0.11	0.98
CI 5					1	0.05
CI 6						1

Table 1: Correlation Coefficients for the Six CI Used in the Study

This correlation between CIs implies that for a given function of distributions to have a threshold that operationally meets the design PFA, the CIs must be whitened (e.g. de-correlated). Fukinaga (1990) presents a whitening transform using the Eigenvector matrix multiplied by the square root for the Eigenvalues (diagonal matrix) of the covariance of the CIs.

$$\mathbf{A} = \Lambda^{1/2}\Phi^T \qquad (3)$$

where ΦT is the transpose of the eigenvector matrix, and Λ is the eigenvalue matrix. The transform can be shown to not be orthonormal, illustrating that the Euclidean distances are not preserved in the transform. While ideal for maximizing the distance (separation) between classes (such as in a Baysian classifier), the distribution of the original CI is not preserved. This property of the transform makes it inappropriate for threshold setting.

If the CIs represented a metric such as shaft order acceleration, then one can construct an HI which is the square of the normalized power (e.g. square root of the acceleration squared). This can be defined as normalized energy, where the health index is:

$$HI = \sqrt{CI \times \operatorname{cov}(CI)^{-1} \times CI^T} \qquad (4)$$

Bechhoefer and Bernhard (2007) were able to whiten the CI and establish a threshold for a given PFA.

The objective of this analysis is to broaden the diagnostic capability available for gear health indexes by generalizing a method to develop HIs across CIs with other functions and statistical distributions.

1. GENERALIZED FUNCTION OF DISTRIBUTIONS

The desired linear transformation operates on the vector CI such that:

$$\mathbf{Y} = \mathbf{L} \times CI^T, \quad 0 = \rho = correlation(\mathbf{Y}) \qquad (5)$$

where Y preserves the original distribution of the CIs.

The Cholesky Decomposition of Hermitian, positive definite matrix results in A = LL*, where L is a lower triangular, and L^* is its conjugate transpose. By definition, the inverse covariance is positive definite Hermitian. It then follows that:

$$\mathbf{LL^*} = \Sigma^{-1} \qquad (6)$$

and

$$\mathbf{Y} = \mathbf{L} \times CI^T \qquad (7)$$

where **Y** is 1 to n independent CI with unit variance (one CI representing the trivial case). The Cholesky Decomposition, in effect, creates the square root of the inverse covariance. This in turn is analogous to dividing the CI by its standard deviation (the trivial case of one CI). In turn, Eq. (7) creates the necessary independent and identical distributions required to calculate the critical values for a function of distributions.

1.1 Gear Health as a Function of Distributions

Prior to detailing the mathematical methods used to develop the HI, background information will be discussed. A common nomenclature for the user/operator of the condition monitoring system will be presented, such that the health index (HI) has a common meaning. The critical values (threshold) will be different for each monitored component, because the measured CI statistics (e.g. covariance) will be unique for each component type. The threshold will be normalized, such that the HI is independent of the component. Further, using guidance from GL Renewables (2007), the HI will be designed such that there are two alert levels: warning and alarm. Then the HI is defined such that the range is:

- 0 to 1, where the probability of exceeding an HI of 0.5 is the PFA
- A warning alert is generated when the HI is greater than or equal to 0.75
- An alarm alert is generated when the HI is greater than or equal to 1.0

2. HI BASED ON RAYLEIGH PDFs

The PDF for the Rayleigh distribution uses a single parameter, β, resulting in the mean ($\mu = \beta*(\pi/2)^{0.5}$) and variance ($\sigma^2 = (2 - \pi/2) * \beta^2$) being a function of β. Note that when applying Eq. (7) to normalize and whiten the vector of CI data, the value for β for each CI will then be:

$$\sigma^2 = 1,$$
$$\beta = \sigma^2 / \sqrt{2 - \pi/2} = 1.5264 \tag{8}$$

The PDF of the Rayliegh is:

$$f(x) = x/\beta^2 \exp\left(-x/2\beta^2\right) \tag{9}$$

The cumulative distribution function, the integral of (9) is:

$$F(x) = 1 - \exp\left(-x^2/2\beta^2\right) \tag{10}$$

It can be shown that the PDF of the magnitude of a frequency of a random signal is a Rayleigh PDF (Bechhoefer and Bernhard 2006). This property makes the Rayleigh an appropriate model for thresholds for shaft (Shaft order 1, etc) and bearing energies. The next section will demonstrate how this can be used appropriately for gears.

2.1 The Rayleigh Order Statistic

Consider a HI function which takes the maximum of n CIs. If the CIs are Independent and Identical (IID), then the function defines the order statistic. Given the order statistic PDF as (Wackerly 1996):

$$g(x) = n\left[F(x)\right]^{n-1} f(n) \tag{11}$$

The threshold is then calculated for t, from the inverse Cumulative distribution function (CDF):

$$1 - PFA = \int_{x=-\infty}^{t} n\left[F(x)\right]^{n-1} f(n)\,dx \tag{12}$$

For n = 3, PFA of 10-3, after solving the inverse CDF (Eq 12), the threshold t equals 6.1 (Note, the solution to Eq 12 can sometime require significant effort. See the Appendix for solution strategies). The HI algorithm, referred to as the Rayleigh Order Statistics (OS) is then:

$$HI = \max\{\mathbf{Y}\} \times {}^{0.5}\!/_{6.1} \tag{13}$$

Here Y = L x CIT (e.g. whitening and normalizing the CIs by applying Eq (7), which is scaled by 0.5 over the threshold. This then is consistent with the definition of the HI presented in 2.1, or a HI of 0.5 for the defined PFA.

2.2 The Sum of *n* Rayleigh

Consider a HI function which takes the sum of n CIs. If the CIs are Independent and Identical (IID), then the function defines a distribution with a Nakagami PDF (Bechhoefer and Bernhard 2007). Given the mean and variance for the Rayleigh, the sum of n normalized Rayleigh distributions is $n * \beta*(\pi/2)^{0.5}$, with variance σ2 = n. Given the Nakagami PDF as:

$$2\left({}^{\eta}\!/_{\omega}\right)^{\eta} \, {}^{1}\!/_{\Gamma(\eta)} \, x^{(2\eta-1)} e^{-{}^{\eta}\!/_{\omega} x^2} \tag{14}$$

where Γ is the gamma function. Then, the statistics for the Nakagami are calculated as:

$$\eta = E\left[x^2\right]^2 / Var\left[x^2\right],\ \omega = E\left[x^2\right] \tag{15}$$

which are used in the inverse Cumulative distribution function (CDF) to calculate the threshold.

For n = 3 CIs, the threshold is 10.125 and the HI algorithm, referred to as the Rayleigh normalized energy (NE) is then:

$$HI = {}^{0.5}\!/_{10.125} \sum_{i=1}^{3} \mathbf{Y}_i \tag{16}$$

For a more in depth treatment of the Nakagami, see Bechhoefer and Bernhard (2007). Again, the dividing 0.5/10.125 allows Eq (15) to be consistent with the HI paradigm.

2.3 The Total Energy of *n* Rayleigh

Consider a HI function which takes the norm of n CIs, which represents the normalized energy. If the CIs are IID, it can be shown that the function defines a Nakagami PDF (Bechhoefer and Bernhard 2007). The mean is now $2*n*1/(2-\pi 2)^{0.5}$. Then, the statistics for the Nakagami are calculated as:

$$\eta = n,\ \ \omega = 1/\left(2 - \pi/2\right) * 2 * n \tag{17}$$

which are used in the inverse CDF to calculate the threshold. For our n = 3 CIs, the threshold is 6.259 and the HI algorithm, referred to as the Sum of Raleigh (SR) is then:

$$HI = {}^{0.5}\!/_{6.259} \sqrt{\sum_{i=1}^{3} \mathbf{Y}_i^2} \tag{18}$$

3. HI BASED ON GAUSSIAN PDFs

If it is found that the distribution of the CI data follows a Gaussian distribution a comparable mathematical process can be applied. Using similar constructs as applied to the Rayleigh PDF, we can generate thresholds for the Gaussian distribution. The PDF of the Gaussian is:

$$f(x) = x/\sigma\sqrt{2\pi}\,\exp\left(-(x-\mu)^2/2\sigma^2\right) \tag{19}$$

The cumulative distribution function, the integral of Eq (19) is

$$F(x) = x/\sigma\sqrt{2\pi}\int_{-\infty}^{x}\exp\left(-(t-\mu)^2/2\sigma^2\right)dt \tag{20}$$

3.1 The Gaussian Order Statistic

Eq. 11 can be applied to the Gaussian PDF and CDF to derive the order statistic PDF of the Gaussian HI function:

$$f(x) = 3\left[x/\sigma\sqrt{2\pi}\int_{-\infty}^{x}\exp\left(-(t-\mu)^2/2\sigma^2\right)dt\right]^2 \times\; x/\sigma\sqrt{2\pi}\,\exp\left(-(x-\mu)^2/2\sigma^2\right) \tag{21}$$

Again, we find the threshold by solving the inverse CDF of Eq (12). The PDF of the order statistic (OS) for a zero mean Gaussian is not bounded at zero, such as the Rayleigh. As such, to be consistent without the HI paradigm of lower HI range of 0, the OS PDF is shifted such the probability of the HI being less than or equal to zero is small. In this example, that probability is defined at 0.05%, corresponding to a PFA of 0.95 (e.g. a lower threshold). For n = 3, for a PFA of 0.95, lower threshold, t is -0.335, and upper threshold for a PFA of 10^{-3}, the threshold t is 3.41 (for HI of 0.5). The CIs are now a z distribution (Gaussian normalized with zero mean and unit variance). An additional rule is set such that any HI less than the lower 5% (corresponding to a PFA of 0.95) is an HI of zero. The HI algorithm is:

$$\begin{aligned}\mathbf{Y} &= \mathbf{L}\times\left(CI^T - \mathbf{m}\right)\\ HI &= \left(\max\{\mathbf{Y}\} + .34\right)\times 0.5/(3.41 + 0.34)\end{aligned} \tag{22}$$

where **m** is the mean of the CIs. Subtracting the mean and multiplying by **L** transforms the CIs into n, Z distributions (zero mean, IID Gaussian distributions).

3.2 The Sum of *n* Gaussian

Consider a HI function that takes the sum of n Gaussian CIs. Then the mean and variance of the sum of the CI are:

$$\mu = \sum_{i=1}^{3} E\left[\mathbf{L}_i\right], \quad \sigma^2 = n \tag{23}$$

Again the inverse normal CDF is used to calculate the threshold. Similar to (22), an offset and scale value is needed to ensure the HI is lower bounded to 0. For n = 3 CI, the mean, μ = 3 and variance σ^2 = 3. Using the inverse normal CDF, the lower threshold (PFA of .95) is -0.15 and the and upper threshold (PFA 10^{-3}), is 8.352, then the HI algorithm is then:

$$\begin{aligned}\mathbf{Y} &= \mathbf{L}\times CI^T\\ HI &= 0.5/(8.352 - 0.15)\left(-0.15 + \sum_{i=1}^{3}\mathbf{Y}_i\right)\end{aligned} \tag{24}$$

3.3 The Total Energy of *n* Gaussian

Finally, we will consider a HI function that takes the norm of n Gaussian CIs. Again it can be shown that the function defines a Nakagami PDF (Bechhoefer and Bernhard 2007). The mean is 2*n* 1/sqrt(2-π2), with ω = n and η is ω/2. Using the inverse Nakagami CDF to calculate the threshold for n = 3 CIs and a PFA of 10^{-3}, the threshold is: 3.368. The HI algorithm is then:

$$\begin{aligned}\mathbf{Y} &= \mathbf{L}\times CI^T\\ HI &= 0.5/3.368\sqrt{\sum_{i=1}^{3}\mathbf{Y}_i^2}\end{aligned} \tag{25}$$

4. APPLICATION TO GEAR FAULT

Vibration data from experiments performed in the Spiral Bevel Gear Test facility at NASA Glenn was reprocessed for this analysis. A description of the test rig and test procedure is given in Dempsey *et al*. (2002). The rig is used to quantify the performance of gear material, gear tooth design and lubrication additives on the fatigue strength of gears. During this testing, CIs and oil debris monitoring were used to detect pitting damage on spiral bevel gears (**Figure 1 Test Rig and Gears (Dempsey et al. 2002)**.

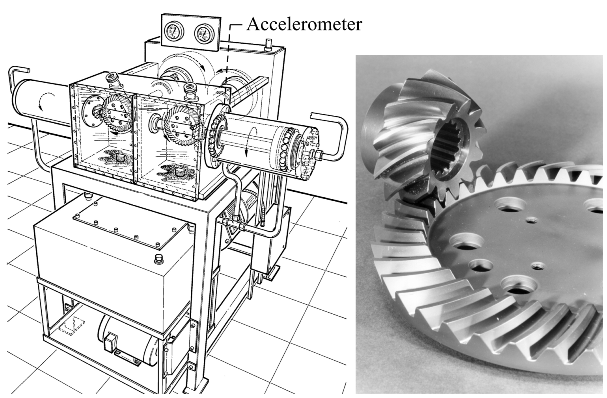

Figure 1 Test Rig and Gears (Dempsey et al. 2002)

The tests consisted of running the gears under load through a “back to back” configuration, with acquisitions made at 1 minute intervals, generating time synchronous averages (TSA) on the gear shaft (36 teeth). The pinion, on which the damage occurred, has 12 teeth.

TSA data was re-processed with gear CI algorithms presented in Zakrajsek *et al*. (1993) and Wemhoff *et al*. (2007), to include:

- TSA: RMS, Kurtosis (KT), Peak-to-Peak (P2P), Crest Factor (CF)
- Residual RMS, KT, P2P, CF
- Energy Operator RMS, KT
- Energy Ratio
- FM0
- Sideband Level factor
- Narrowband (NB) RMS, KT, CF
- Amplitude Modulation (AM) RMS, KT
- Derivative AM KT
- Frequency Modulation (FM) RMS, KT

From these CIs, a total of six CIs were used for the HI calculation: Residual RMS, Energy Operator RMS, FM0, NB KT, AM KT and FM RMS. These CIs were chosen because they exhibited good sensitivity to the fault. Residual Kurtosis and Energy Ratio also were good indicators, but were not chosen because;

- It has been the researcher's experience that these CIs become ineffective when used in complex gear boxes, and
- As the faults progresses, these CIs lose effectiveness. The residual kurtosis can in fact decrease, while the energy ratio will approach 1.

Covariance and mean values for the six CI were calculated by sampling healthy data from four gears prior to the fault propagating. This was done by randomly selecting 100 data points from each gear, and calculating the covariance and means over the resulting 400 data points.

The selected CI's PDF were not Gaussian, but exhibited a high degree of skewness. Because of this, the PDFs were "left shifted" by subtracting an offset such that the PDFs exhibited Rayleigh like distributions. Then, the threshold setting algorithms were tested for:

- Rayleigh order statistic (OS): threshold 8.37 for n = 6 and a PFA of 10^{-6},
- Rayleigh normalized energy (NE): threshold 10.88 for n = 6 and a PFA of 10^{-6},
- Sum of Rayleigh (SR): threshold 24.96 for n = 6 and a PFA of 10^{-6},

Figures 2, 4 and 6 are HI plots that compare the OS, NE and SR algorithms during three experiments in the test rig. The HI trend (in black) is plotted on top of the raw HI values (in blue). Figures 3, 5 and 7 show the amount of pitting damage on the pinion teeth at each test completion.

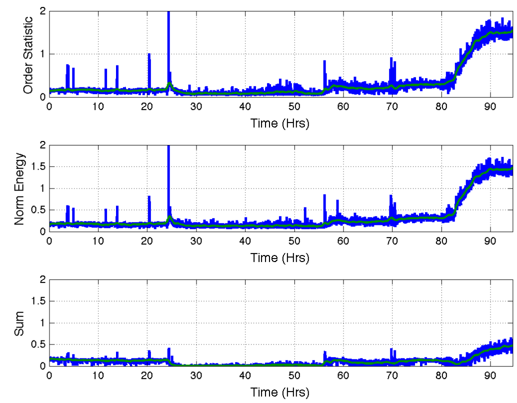

Figure 2 Test BV2_10_15_01EX4

Note that the spikes corresponded to changes in torque on the rig. All the HI algorithms where sensitive to damage, although in general, the best system response was from both the OS and NE.

Figure 3 Pitting Damage on EX4

Note that the decrease in the HI rate of change corresponds to a decrease in torque load towards the end of the test.

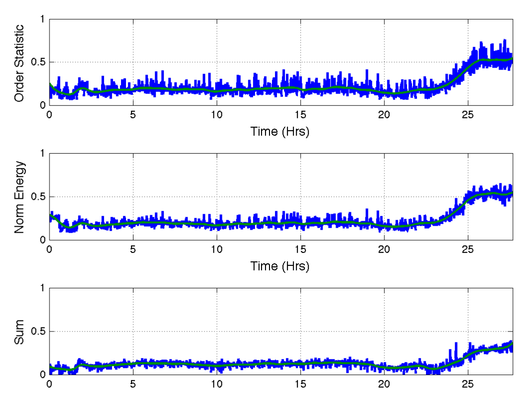

Figure 4 Test BV2_10_26_01EX5

For the data plotted in figure 4, this test appears to have been halted prior to heavy pitting damage, as the gear HI is reach only 0.5. However, the photo of gear EX5 (Figure 5) shows extensive pitting damage.

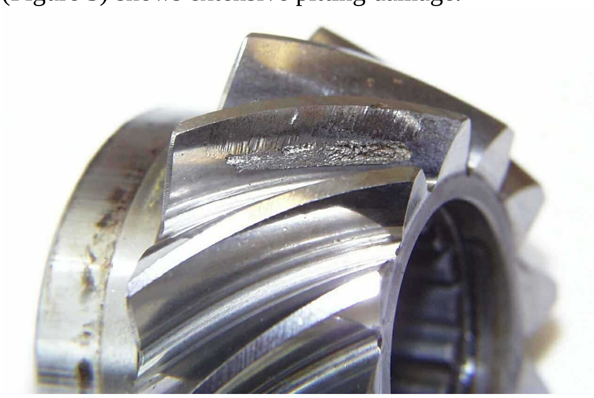

Figure 5 Damage on Gear EX5

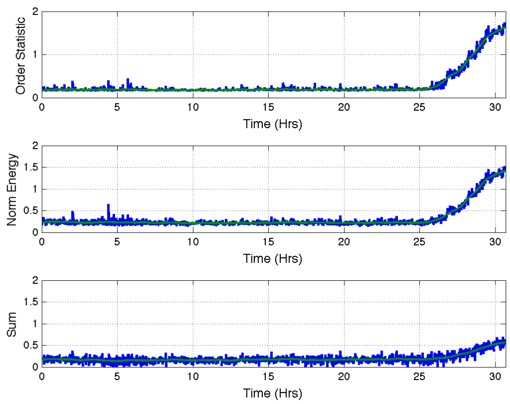

Figure 6 Test BV2_1_4_02EX6

Figure 7 Damage on Gear EX6

5. DISSCUSION AND OBSERVATIONS

After the three statistical models were applied to the test rig CI data, it was observed that each HI algorithm performed well, although the OS and NE is clearly more sensitive to fault than the SR algorithm. Additionally, the measured RMS noise of the OS was 15% to 25% higher than the NE, that RMS value being approximately 0.05 HI. However, the most important contribution is that a process has been developed to whiten CI data so that different HI algorithms can be explored with some assurance that, mathematically, the PFA performance was being met.

Additionally, it is encouraging that, based solely on nominal data (statistics taken prior to fault propagation), it was observed that:

- An HI of 1 displays damage warranting maintenance.
- That nominal data is approximately 0.1 to 0.2 HI, where the PFA was set for 0.5 HI
- That while no one CI seemed to work for every gear tested, the HI function captured the damage consistently (even for a small sample set).
- The HI trends were low noise. This can facilitate prognostics.

6. CONCLUSION

Thresholding is critical for the operation of a condition monitoring system. If the probability of false alarm (PFA) is too high, then the operator is flooded with numerous false alarms and tends to disregards alerts. Unfortunately, some of the alerts will be true, resulting in collateral equipment damage. If the PFA is low, but the probability of fault detection is low, then the operator cannot perform maintenance “on condition”. Again, there are missed faults resulting in collateral damage.

Because the condition indicators (CI) are correlated, without some pre-processing, it is difficult to operationally achieve the design PFA. A method was presented for whitening the CIs used in gear fault detection. The whitening was achieved by a linear transformation of the CI using the Cholesky decomposition of the inverse of the CIs covariance.

With this transformed, whitened CI data, a health indexed based on a specified PFA was demonstrated. Three candidate HI algorithms (order statistics, normalized energy and sum of CI) for two different CI probability distribution functions (Gaussian and Rayleigh), were presented and tested on three data sets of pitted gears from a test stand.

It was observed that the HI algorithms performed as designed: low PFA (e.g. noise) and good fault detection capability. Given this process, we will now expand the class of distributions that this can be applied to, for example, the Rice and Weibull distribution.

APPENDIX: Monte Carlo Techniques to Solve the Inverse CDF

The solution of the inverse CDF can be difficult for none standard distribution. In fact, most function of distributions are non-standard. Solutions for order statistic on Gaussians distribution are very problematic:

even solving using optimization techniques is nontrivial.

Alternatively, Monte Carlo techniques are relatively simple to set up, and give accuracy limited only by patients. For example, since the order statistic is defined as the maximum of n IID distribution, it is relatively easy to call 10 million random tuples of n distribution, take the maximum of each tuple, and sort to generate the CDF. The critical value corresponds to the index of the sorted values at 10 million x (1-PFA).

As an experiment, find the inverse CDF for the normal Gaussian with a PFA of 10-3. For 10 million, the index is 9990000. Running 100 experiments, the estimated critical value was: 3.090155199948529 vs. the actual value of 3.090232306167824. The PFA calculate from the Monte Carlo generated threshold was: 0.00100025, or an error of .025%.

REFERENCES

McFadden, P., Smith, J., (1985), *A Signal Processing Technique for detecting local defects in a gear from a signal average of the vibration.* Proc Instn Mech Engrs Vol 199 No C4

Zakrajsek, J. Townsend, D., Decker, H., (1993). *An Analysis of Gear Fault Detection Method as Applied to Pitting Fatigue Failure Damage*. NASA Technical Memorandum 105950.

Lewicki, D., Dempsey, P., Heath, G., and Shanthakumaran P. (2010), *Gear Fault Detection Effectiveness as Applied to Tooth Surface Pitting Fatigue Damage*, Gear Technology, November/December 2010.

Wemhoff, E., Chin, H., Begin, M., (2007), *Gearbox Diagnostics Development Using Dynamic Modeling*, AHS 63rd Annual Forum, Virginia Beach, 2007

Byington, C., Safa-Bakhsh, R., Watson., M., Kalgren, P., (2003), *Metrics Evaluation and Tool Development for Health and Usage Monitoring System Technology*, HUMS 2003 Conference, DSTO-GD-0348

Wackerly, D., Mendenhall, W., Scheaffer, R.,(1996), Mathematical Statistics with Applications, Buxbury Press, Belmont, 1996

Fukunaga, K., (1990), Introduction to Statistical Pattern Recognition, Academic Press, London, 1990, page 75.

Bechhoefer, E., Bernhard, A., (2007), *A Generalized Process for Optimal Threshold Setting in HUMS*, IEEE Aerospace Conference, Big Sky.

GL Renewables, (2007), *Guidelines for the Certification of Condition Monitoring Systems for Wind Turbines*, http://www.gl-group.com/en/certification/renewables/CertificationGuidelines.php

Bechhoefer, E., Bernhard, A., (2006), *Use of Non-Gaussian Distribution for Analysis of Shaft Components*, IEEE Aerospace Conference, Big Sky.

Dempsey, P., Handschuh, R., Afjeh, A. (2002), *Spiral Bevel Gear Damage Detection Using Decision Fusion Analysis*, NASA/TM-2002-211814