ETD Archive

2011

Quality Control Using Inferential Statistics in Weibull Analyses for Components Fabricated from Monolithic Ceramics

Ankurben H. Parikh
Cleveland State University

QUALITY CONTROL USING INFERENTIAL STATISTICS IN WEIBULL

ANALYSES FOR COMPONENTS FABRICATED FROM MONOLITHIC

CERAMICS

ANKURBEN PARIKH

Bachelor of Engineering in Civil - Irrigation and Water Management

The Maharaja Sayajirao University

May 2003

Master of Engineering in Civil – Water Resources Engineering

The Maharaja Sayajirao University

May 2005

Submitted in partial fulfillment of requirements for the degree

MASTER OF SCIENCE IN CIVIL ENGINEERING

at the

CLEVELAND STATE UNIVERSITY

FALL 2011

This thesis has been approved

for the Department of Civil Engineering

and the College of Graduate Studies by

_____

Thesis Chairperson, Dr. Stephen F. Duffy

_____

Department & Date

_____

Dr. Lutful I. Khan

_____

Department & Date

_____

Dr. Jacqueline Jenkins

_____

Department & Date

ACKNOWLEDGEMENTS

QUALITY CONTROL USING INFERENTIAL STATISTICS IN WEIBULL BASED

ANALYSES OF COMPONENTS FABRICATED FROM MONOLITHIC CERAMICS

ANKURBEN PARIKH

ABSTRACT

This work presents the mathematical constructs for certain statistical elements that when combined properly produce a quality control program that can be used to accept or reject ceramic materials based on mechanistic strength information. Due to the high tensile strength and low fracture toughness of ceramic materials the design engineer must consider a stochastic design approach. Critical flaws with lengths that cannot be detected by current non-destructive evaluation methods render a distribution of defects in ceramics that effectively requires that the tensile strength of the material must be treated as a random variable. The two parameter Weibull distribution (an extreme value distribution) with size scaling is adopted for tensile strength in this work.

Typically the associated Weibull distribution parameters are characterized through the use of four-point flexure tests. The failure data from these tests are used to determine the Weibull modulus (m) and a Weibull characteristic strength (σ_θ). To determine an estimate of the true Weibull distribution parameters maximum likelihood estimators are used. The quality of the estimated parameters relative to the true distribution parameters depends fundamentally on the number of samples taken to failure and the component under design. The statistical concepts of "confidence intervals" and "hypothesis testing" are discussed here relative to their use in assessing the "goodness" of the estimated distribution parameters. Both of these inferential statistics tools enable the

calculation of likelihood confidence rings. Work showing how the true distribution parameters lie within a likelihood ring with a specified confidence is presented.

A material acceptance criterion is defined here and the criterion depends on establishing an acceptable probability of failure of the component under design as well as an acceptable level of confidence associated with estimated distribution parameter determined using failure data from a specific type of strength test. A standard four point bend bar was adopted although the procedure is not limited to only this type of specimen. This work shows how to construct likelihood ratio confidence rings that establishes an acceptance region for distribution parameters relative to a material performance curve. Combining the two elements, i.e., the material performance curve based on an acceptable component probability of failure and a likelihood ratio ring based on an acceptable confidence level, allows the design engineer to determine if a material is suited for the component design at hand – a simple approach to a quality assurance criterion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

A = Area of test specimen

A_E = Effective surface parameter

A_{comp} = Failure surface of the component

$A_{effective}$ = Effective area of the four point bend bar

A_{inner} = Inner area of the thin wall pressure vessel

A_{outer} = Outer area of the thin wall pressure vessel

F = Load at the fracture point

H_0 = Null hypothesis

H_1 = Alternative hypothesis

K_{IC} = Fracture toughness

L = Likelihood function

L_i = Length of the inner load span

L_o = Length of the outer load span

N = Number of sample size

P_f = Probability of failure

$(P_f)_{component}$ = Probability of failure of the component

R = Reliability of the continuum element

R_i = Reliability of the ith sub-element

T = Test statistic

Y = Geometry factor

b, d = Width and Depth of the specimen

c	=	Half the crack length
m	=	Weibull modulus
m_A	=	Weibull modulus associated with surface area
$\widetilde{m}$	=	Estimated Weibull parameter
p	=	Applied internal pressure in thin wall pressure vessel
r	=	Radius of thin wall pressure vessel
r_{avg}	=	Average radius to the middle of the pressure vessel wall
t	=	Wall thickness of thin wall pressure vessel
α	=	Significance level
β	=	Scale parameter
γ	=	Confidence level
σ	=	Applied far field stress
σ_θ	=	Weibull characteristic strength
$(\sigma_\theta)_A$	=	Weibull characteristic strength associated with surface area
σ_0	=	Weibull material scale parameter
$(\sigma_0)_A$	=	Weibull material scale parameter associated with surface area
σ_{max}	=	Maximum tensile stress in a test specimen at failure
$\widetilde{\sigma_\theta}$	=	Estimated Weibull characteristic strength
Θ_0	=	Contains all the MLE parameter estimates
Θ_0^C	=	All point estimates that are not MLE parameter estimates
$\mathcal{L}$	=	Natural log of the likelihood function

CHAPTER I

INTRODUCTION

1.1 Introduction

Ceramic materials possess numerous advantageous properties such as high strength at elevated use temperatures, low density, high stiffness, superior wear properties and inert behavior in aggressive environments. Ceramics can be thermally conductive or act as insulators. They can be electronically opaque or translucent. These attributes make ceramics attractive in engineering applications. There are many potential applications that can take the advantage of the strong mechanical, thermal and electrical properties of this material. When compared to high temperature conventional ductile materials, advanced ceramics possess high-strength properties at elevated service temperatures. Ceramics have demonstrated functional abilities at temperatures above 1300 °C. This is well beyond the operational limits of most conventional metal alloys used in advanced diesel and turbine engine applications.

Specific high value added applications of ceramics include rocket nozzles, hip and knee prosthetics, dental crowns and dental bridges. This material is used to fabricate industrial parts such as cutting tools, grinding wheels and bearings. Ceramics are used in

microturbines, pressure sensors and thin film membranes in micro-electro-mechanical systems (MEMS). Oxygen ion-conducting ceramic membranes are used in solid oxide fuel cells, extra-corporeal membrane oxygen (ECMO) separators and catalytic membrane reactors. Military applications of ceramics include missile radomes (electronically opaque), advanced armor, blast protection systems for tactical wheeled-vehicle fleet and advanced gun barrel systems. The range of applications for ceramics is quite broad.

Even though ceramics are extremely useful in high temperature and/or wear resistant applications, designing components fabricated with this material is not straight forward. Ceramic materials exhibit relatively high tensile strength with a comparatively low fracture toughness (quantified by K_{IC}). Lack of fracture toughness leads to low strain tolerance and huge variations in fracture strength. Current nondestructive evaluation (NDE) methods cannot detect small critical defect sizes. The minute critical defect size is a result of the combination of high strength and low fracture toughness. The random distributions of these defects produce the observed scatter in tensile strength associated with ceramics. These defects have various crack sizes and orientations. If the distribution of the size and orientation of defects is random and distributed homogeneously, then this leads to a decrease in tensile strength as the size of the specimen increases. Due to this strength-size effect, bigger components have a higher probability of deleterious flaws present. Design methods for ceramics must accommodate this strength-size effect. This is done using system reliability concepts. The ceramic component is treated as a system and the probability of system failure is ascertained. Due to inherent scatter in strength and size effects, reliability analysis of components fabricated from advanced ceramics is preferred over the use of standard

safety factors for acceptable loading. Statistical analysis should be an integral part of the design procedure and this requires accepting a finite risk of failure.

Typically, finite element methods are used for designing brittle material structural components subjected to complex boundary conditions. A number of reliability algorithms such as CARES (1990) and STAU (1991) have been coupled with commercial finite element software. These reliability algorithms allow for component failure to originate in any discrete element. This is a result of considering the component as a system and using system reliability theories. When one discrete element of series system fails, then a component is modeled by weakest link reliability theories. However, in parallel systems, when one element fails the remaining elements maintain the applied load through redistribution. This type of failure is characterized through the use of bundle theories and tends to be used in the design of ceramic composites which won't be considered here. In reliability analysis, weakest-link theories and bundle theories represent the bounds on the probability of system failure. One can also consider more complex systems such as "r out of n" systems where at least r elements have to fail before the system fails. This produces intermediate values of the probability of failure relative to weakest link and bundle theories. Returning the focus to monolithic ceramics, these materials fail in a sudden and catastrophic manner. So a weakest-link, i.e., a series system is adopted.

Ceramics have relatively low K_{IC} values, i.e., in the range of 1-12 MPa $\cdot$ m$^{1/2}$. When combined with high tensile strengths, this produces a distribution of flaw sizes assuming mode I failure, characterized by the following simple relationship:

$$\sigma \;=\; \frac{K_{IC}}{Y(\pi c)^{\frac{1}{2}}} \tag{1.1}$$

Here, σ is the applied far field stress, Y is a geometry factor and c is half the crack length. Based on low K_{IC} values, critical flaw sizes produced in the fabrication process are extremely small (well beyond detection limits), so a random distribution of flaw sizes and orientations occur and results in a wide variation of tensile strengths. Due to this variation, tensile strength must be considered as a random variable. Failure strength is taken either for a comparison of the relative quality of two materials, or for the prediction of the probability of failure when designing a component fabricated from ceramic materials. The calculation of a component probability of failure first requires characterizing parameters for the distribution function adopted to describe the tensile strength random variable. Characterizing distribution parameters is accomplished via strength tests; the most common are flexure tests and tensile tests. Of the two, tensile tests are preferred if they can be performed properly.

Tensile tests are useful for quality control purposes and also to predict how a material fails under a uniform stress state. This specimen geometry tends to interrogate flaw populations distributed through the volume of the specimen. Alignment of tensile test specimens is very critical for brittle materials, since misalignment produces bending stresses that affect the failure stress associated with the specimen. This situation can be managed with careful attention to alignment procedures.

Flexural strength is the measure of the ultimate strength of a beam test specimen. The bend test is conducted using specimen geometry with either a rectangular or circular cross section. Since the maximum tensile stress occurs at the outer surface of this

specimen and a linear stress gradient is present from the neutral axis to the outer surface, this specimen geometry tends to interrogate flaws along the surface and subsequently assesses the material's surface finish. The specimen is failed using either a four point or three point load configurations. Four point flexure testing specimen is symmetrically loaded at two locations that are typically situated at the quarter points. This test introduces a uniaxial state of stress within the gage section (between the inner load points) of specimen. For a rectangular specimen under a load in four-point bending, stress is expressed as

$$\sigma = \frac{3F(L_o - L_i)}{2bd^2} \tag{1.2}$$

Where,

F =   load at the fracture point

L_i =   length of the inner load span

L_o =   length of the outer load span

b, d =   width and depth of the specimen

The three point flexure test specimen is an alternative specimen geometry that produces lower quality data. Here the specimen is loaded at a location midway between two support bearings. This test introduces shear stresses between the load points and is considered inferior test specimen geometry when interrogating the tensile strength of the material. For a rectangle specimen under a load in a three-point bending specimen, stress is given by

$$\sigma = \frac{3FL}{2bd^2} \tag{1.3}$$

For all uniaxial tensile specimens we assume that the scatter in ceramic tensile strength is characterized by the two parameter Weibull distribution. The probability that the random variable X assumes a value between a and b is given by the expression

$$P_r(a < X < b) \quad = \quad \int_b^a f(x)dx \qquad (1.4)$$

The continuous random variable X has a two-parameter Weibull distribution if the probability density function is given by

$$f(x) \quad = \quad \left(\frac{m}{\beta}\right)\left(\frac{x}{\beta}\right)^{m-1} \exp\left[-\left(\frac{x}{\beta}\right)^m\right] \qquad x > 0 \qquad (1.5)$$

$$f(x) \quad = \quad 0 \qquad x \le 0 \qquad (1.6)$$

and the cumulative distribution function is given by

$$F(x) \quad = \quad 1 \quad - \quad \exp\left[-\left(\frac{x}{\beta}\right)^m\right] \qquad x > 0 \qquad (1.7)$$

$$f(x) \quad = \quad 0 \qquad x \le 0 \qquad (1.8)$$

Here, m is the shape parameter or "Weibull modulus" *(>0)* and β is known as scale parameter *(>0)*. For ceramic strength tests the industry refers to β as a "characteristic strength" and is designated as σ_θ. Both parameters are determined through the analysis of tensile strength data from the specimens geometries just discussed.

There are at least three methods that can be utilized in parameter estimation, including the method of moments, linear regression techniques and maximum likelihood techniques. These approaches yield optimized estimates based on specified criterion. At times the estimators from all three techniques coincide, most times they do not. Here we will always estimate Weibull parameters using maximum likelihood estimator (MLE).

Estimates are approximate values of the true parameter values. The method of maximum likelihood is a commonly used estimation technique because this type of estimator maintains some attractive features. Specifically, these estimates are efficient in that they produce the least amount of bias and they are invariant as well.

Weibull distribution parameters are estimated on the basis of observations from a number of strength tests. Usually there is no preconceived notion regarding the value of the Weibull distribution parameters when sampling and testing, but obviously there is a preconception regarding the type of distribution that characterizes the population. The Weibull distribution is an extreme value distribution and this facet makes it a preferred distribution to characterize material "strength." We wish to characterize a material's minimum strength. Thus the first step is simply ascertaining values for the distribution parameters. The next issue is have we estimated parameters to the best of our ability? This is directly related to the fundamental question asked repeatedly "how many samples must be tested?" The typical ceramics industry answer to this question is 30. However, the question that should be asked is "how many samples must be tested to establish a given confidence level for component reliability?" Helping to answer this question quantitatively is the goal of this thesis. To address the issue properly one needs to utilize interval estimates along with hypothesis testing.

Confidence intervals are used to indicate the reliability of estimated parameters. Keep in mind that parameter estimates are random variables. Every time a sample is taken from the same population a point estimate is made using maximum likelihood estimators and rarely do successive samples produce the same point estimate values. Thus interval estimates are as necessary as point estimates. Interval estimates will reflect

our confidence that the true population parameters actually belong to their respective intervals, as well as give an indication of the amount of variability that is associated with the point estimates. A large confidence interval indicates that the parameter estimates maintain a high degree of uncertainty. Increasing the sample size will always narrow confidence intervals unless there is a deficiency with the point estimation functions. In essence, confidence intervals on parameter estimates represent the range of values for the distribution parameters that are both reasonable and plausible.

In contrast, when testing a hypothesis on distribution parameters, there is a preconceived notion of the value the distribution parameters take on. There are two statistical hypotheses proposed concerning the estimated parameter values. The hypotheses are then tested with the samples (data) taken from the distribution populations. The two hypotheses are the one being proposed by the analyst, and the negation of this hypothesis. The former, denoted H_1, is referred to as the alternative hypothesis, but more accurately this should be called the analyst's hypothesis. The latter hypothesis, denoted here as H_0, is referred to as the null hypothesis. The goal of the analyst is to decide whether there is enough evidence (data) to refute the null hypothesis. The decision is made by observing the value of a well defined test statistic which is ratio of two likelihood functions whose probability is known under the assumption that H_0, the null hypothesis, is true. This statistic lends credence to the alternative hypothesis, H_1, if the statistic takes on a value rarely encountered using the data collected. If this happens to an unacceptable degree then the null hypothesis is rejected. The value of the test statistic at which the rejection is made defines a rejection region. The probability that the

test statistic falls into the rejection region by chance is referred to as the significance level, and is denoted α.

1.2 Scope

We wish to apply the concepts outlined above for the specific case where the tensile strength of a ceramic material can be characterized by the Weibull distribution, which is an extreme value distribution. A two-parameter Weibull distribution with size scaling is used here to characterize the tensile strength and strength data from four-point flexure tests are used to estimate the Weibull distribution parameters. To bring out certain underlying principles of the procedures used here tensile strength is generated from Monte-Carlo simulations. Maximum likelihood estimators are used. The quality of the estimated distribution parameters depends on number of sample size taken. Here confidence intervals and hypothesis testing are used to ascertain the quality of the parameter estimates. Material acceptance and rejection regions are defined and inferential statistics are used to generate likelihood confidence rings. Likelihood confidence rings establish an acceptance region for the estimated distribution parameters. A thin wall pressure vessel example is used as an example of how to develop material performance curves and the corresponding confidence ring.

1.3 Objective

It is assumed that the tensile strength is the paramount design parameter for components fabricated from ceramic or graphite materials. To assess the quality of the material relative to a component application the component fabricator needs a quality control process designed to combine information regarding the reliability of the component as well as data from the material supplier regarding the parameter estimates. The objective of this thesis is to define in detail a process that yields a useful tool for the component fabricator that does just that, i.e., develop a quality control process relative to strength testing of ceramic materials that establishes robust material acceptance criteria. It is presumed that the test data mentioned above will be available to all parties. This quality control process is described in the following chapters.

CHAPTER II

RELIABILITY BASED FAILURE ANALYSIS AND STRENGTH TESTS FOR

COMPONENTS FABRICATED FROM CERAMIC MATERIAL

Advanced ceramics typically exhibit linear stress-strain behavior to failure. Strength is determined based on intrinsic fracture toughness and a distribution of flaws present in the material. Monolithic ceramics fail due to the catastrophic propagation of flaws present in the material - minute flaws that result from high strength and low fracture toughness cannot be detected. Hence tensile strength is a random variable.

We know that a fundamental principle of structural reliability is that as the complexity of a system increases the reliability of the system decreases if the design engineer does not compensate in some fashion. Assume that a differential element is subjected to a general three dimensional state of stress, and that the state of stress is characterized by the principal stresses $(\sigma_1, \sigma_2, \sigma_3)$. Load, as characterized by the applied stresses, is a deterministic variable and also is a function of position, i.e., the principal stresses associated with the state of stress at a point can be characterized as follows:

$$\sigma_1 = \sigma_1(x, y, z) \tag{2.1}$$

$$\sigma_2 = \sigma_2(x, y, z) \tag{2.2}$$

$$\sigma_3 \quad = \quad \sigma_3(x, y, z) \tag{2.3}$$

Note that we will assume that the principal stresses do not interact with each other to produce failure, i.e., the basic tenet of the principle of independent action. Thus failure may occur independently due to any of the tensile principal stresses. Recall that we earlier assumed that a two-parameter Weibull distribution characterizes the tensile strength of the ceramic material. The reliability of a continuum element under the application of any of the three principal stresses is given by

$$R_1 \quad = \quad \exp\left[-\left(\frac{\sigma_1}{\sigma_o}\right)^m\right] \tag{2.4}$$

$$R_2 \quad = \quad \exp\left[-\left(\frac{\sigma_2}{\sigma_o}\right)^m\right] \tag{2.5}$$

$$R_3 \quad = \quad \exp\left[-\left(\frac{\sigma_3}{\sigma_o}\right)^m\right] \tag{2.6}$$

Here m and σ_0 are parameters associated with the assumed Weibull distribution. The parameter m is the Weibull modulus and σ_0 is strength like parameter per unit volume. Assuming that three failure modes are statistically independent, and the reliability of the continuum element is

$$R \quad = \quad R_1 R_2 R_3 \tag{2.7}$$

Substitution leads to

$$R \quad = \quad \exp\left[-\left(\frac{\sigma_1}{\sigma_o}\right)^m\right]\exp\left[-\left(\frac{\sigma_2}{\sigma_o}\right)^m\right]\exp\left[-\left(\frac{\sigma_3}{\sigma_o}\right)^m\right] \tag{2.8}$$

or

$$R = \exp\left\{\left[-\left(\frac{\sigma_1}{\sigma_o}\right)^m\right] + \left[-\left(\frac{\sigma_2}{\sigma_o}\right)^m\right] + \left[-\left(\frac{\sigma_3}{\sigma_o}\right)^m\right]\right\} \qquad (2.9)$$

Hu (1994), showed how the reliability of continuum elements in a structure can be combined to yield the reliability of the entire structure. From Hu's (1994) work the reliability of the structure, or component, is given by

$$R_{component} = \lim_{N \to \infty}\left\{\prod_{i=1}^{N} R_i\right\} \qquad (2.10)$$

Here R_i is the reliability of the ith sub-element.

2.1 Size Scaling

The Weibull distribution is ideal for ceramic materials because it allows for strength scaling. As components, or test specimens, increase in size the probability increases that a larger critical flaw is present and the average strength of the material being tested decreases. Assuming that strength experiments are conducted quickly such that slow crack growth does not occur, then the observed strength value is partially dependent on test specimen size, geometry and stress state also material K_{IC} and flaws. We further assume that failure is an isotropic phenomenon, i.e., there are no preferred orientations of strength. With only one principal tensile stress present in a test specimen then the failure probability of specimen is given by cumulative distribution function,

$$P_f = 1 - \exp\left[-\left(\frac{\sigma_{max}}{\sigma_\theta}\right)^m\right] \qquad \sigma_{max} > 0 \qquad (2.11)$$

and

$$P_f = 0 \qquad \sigma_{max} \le 0 \qquad (2.12)$$

13

Here, P_f is the probability of failure, σ_{max} is the maximum tensile stress in a test specimen at failure, and σ_θ is the Weibull characteristic strength.

As noted before, strength controlling flaws can be distributed through the volume, or over the surface, or along the edges of a test specimen. For volume-distributed flaw populations the characteristic strength is designated $(\sigma_\theta)_V$ and the Weibull modulus is designated m_V. For surface-distributed flaws, the characteristic strength is designated $(\sigma_\theta)_A$ and the Weibull modulus is designated m_A. Also, for edge-distributed flaw populations the characteristic strength is designated $(\sigma_\theta)_L$ and the Weibull modulus is designated m_L. Surface distributed flaw distributions are discussed thoroughly in the next section. Volume and edge flaw distributions follow similar treatments. The issue of multiple flaw populations are (surface, volume, edge) is mentioned for completeness as it relates to how various test specimens and components interrogate certain types of flaw distributions.

2.2 Flaw Distribution Along The Surface

Assuming that failure is caused by a flaw located along the surface of a test specimen then the probability of failure is given by the following surface integration.

$$P_f = 1 - \exp\left[\int\left(\frac{\sigma}{(\sigma_o)_A}\right)^{m_A} dA\right] \qquad \sigma_{max} > 0 \qquad (2.13)$$

$$P_f = 0 \qquad \sigma_{max} \leq 0 \qquad (2.14)$$

In the above equation, integration is performed over regions of the test specimen where the principal stress values are tensile. The Weibull modulus (m_A) and the Weibull material scale parameter $(\sigma_0)_A$ are associated with the strength controlling flaws

distributed along the surface. The Weibull material scale parameter has units of stress·area$^{(1/m_A)}$ so that P_f is dimensionless. The integration defined by the above equation results in the following expression

$$P_f = 1 - \exp\left[-kA\left(\frac{\sigma_{max}}{(\sigma_0)_A}\right)^{m_A}\right] \qquad (2.15)$$

Here k is a dimensionless load factor and σ_{max} is the maximum stress at failure. Therefore,

$$(\sigma_0)_A = (\sigma_\theta)_A (kA)^{\frac{1}{m_A}} \qquad (2.16)$$

Alternatively,

$$(\sigma_0)_A = (\sigma_\theta)_A (A_E)^{\frac{1}{m_A}} \qquad (2.17)$$

Where an effective surface parameter is introduced and defined as

$$A_E = kA \qquad (2.18)$$

We now have a relationship between the Weibull scale parameter $(\sigma_0)_A$, which is a material strength parameter, and the Weibull characteristic strength $(\sigma_\theta)_A$ which is a strength measure always dependent on the test specimen geometry. The effective surface A_E is functionally dependent on m_A and the specimen geometry. Similar expressions can be developed for distributions distributed through the volume of the test specimen (see ASTM C 1239-10), or along the edges of the test specimen (Wereszczak, 2010).

2.3 Maximum Likelihood Estimators

Throughout this effort Monte-Carlo simulation is used to generate "data sets" with the requisite random characteristic strength values. By using well defined functions

that incorporate the failure data one can obtain point estimates of the unknown Weibull parameters. These defined functions are known as estimators. As noted earlier the types of estimators include moment estimators, least-squares estimators and maximum likelihood estimators. Maximum likelihood estimators (MLE) are most useful due to the efficiency and the ease application when censored failure populations are present. Here censored data which has more than one flaw population will not be considered.

The parameter estimates calculated using the maximum likelihood technique, are unique and as the size of the sample increases, the estimates statistically approach the true values of the population parameters. The likelihood function used to derive the estimation function is constructed such that the likelihood function is dependent on the two unknown Weibull distribution parameters m and σ_θ. The likelihood function for the two-parameter Weibull distribution for a single-flaw distribution is given by the following expression (see ASTM Standard C1239-2010)

$$L \;=\; \prod_{i=1}^{N}\left(\frac{\tilde{m}}{\tilde{\sigma}_\theta}\right)\left(\frac{\sigma_i}{\tilde{\sigma}_\theta}\right)^{\tilde{m}-1}\exp\left[-\left(\frac{\sigma_i}{\tilde{\sigma}_\theta}\right)^{\tilde{m}}\right] \tag{2.19}$$

Note here that $\widetilde{m}$ and $\widetilde{\sigma}_\theta$ denote estimates of the Weibull distribution parameters whereas m and σ_θ denote the true population parameters, which are always unknown. A system of equations is obtained by differentiating the log likelihood function with respect to $\widetilde{m}$ and $\widetilde{\sigma}_\theta$ and setting each derivative equal to zero, i.e.,

$$\frac{\displaystyle\sum_{i=1}^{N}(\sigma_i)^{\tilde{m}}\ln(\sigma_i)}{\displaystyle\sum_{i=1}^{N}(\sigma_i)^{\tilde{m}}} \;-\; \frac{1}{N}\sum_{i=1}^{N}\ln(\sigma_i) \;-\; \frac{1}{\tilde{m}} \;=\; 0 \tag{2.20}$$

and

16

$$\tilde{\sigma}_{\theta} \quad = \quad \left[\frac{1}{N} \sum_{i=1}^{N} (\sigma_i)^{\tilde{m}} \right]^{\frac{1}{\tilde{m}}} \tag{2.21}$$

where N is the number of failed specimens in the sample set tested. Equation (2.20) is solved numerically for $\tilde{m}$ which is subsequently used in equation (2.21) to solve for $\tilde{\sigma}_{\theta}$.

CHAPTER III

LIKELIHOOD RATIO CONFIDENCE BOUNDS AND INFERENTIAL STATISTICS

Having obtained samples from Mote-Carlo simulation our tensile strength tests we now would like to draw conclusions about the population we sampled from. We would like to know if our sample was large enough such that the parameters estimated are in the same statistical neighborhood as the true population parameters. The techniques for making this kind of assessment utilize inferential statistics. The type of inference that is focused on here is given that the underlying population can be characterized by a two parameter Weibull distribution this is assumed for actual ceramics, what are the bounds on the true population parameter (both of which are unknown) given a particular sample. In general small bounds are better than large bounds, but bounds on estimates must be objective in what they characterize. Sampling from a population generates information that allows us to calculate point estimates of the true distribution parameters. But they are estimates. We note ahead of the discussion below that the bounds on the true population parameter values are dependent upon the sample size (N) and the data values used to compute the point estimates. The focus here will be on a particular type of bound referred to as "likelihood ratio confidence bounds."

But before that discussion takes place some history is presented regarding goodness-of-fit approaches as well as background on hypothesis testing as it pertains to this effort.

K. Pearson (1895) outlined a goodness-of-fit test that spoke to the applicability of a statistical distribution chosen for a population under study. Initially there was little use for this type of analysis because there were so few alternatives to the normal (Gaussian) distribution available at the time. K. Pearson's work (1900a) pointed to the formulation and use of the χ^2-distribution. Fisher (1922b, 1935) saw the need for hypothesis testing and developed the concept of the null hypothesis. The null hypothesis as proposed by Fisher utilizes a general default position that is never proved or established, but is possibly disproved in the course of experimentation or sampling. Subsequent to Fisher's work the concept of the alternative hypothesis was put forth by Neyman and E. Pearson (1933), and it is used to formulate the Neyman–Pearson lemma. Their lemma holds that when conducting hypothesis testing the power of a test statistic, here it is ratio of two likelihood functions, can be characterized by the likelihood ratio test. This lemma serves as a major component in modern statistical hypothesis testing and the "likelihood ratio test" is used here. However, the alternative hypothesis was not part of Fisher's approach of statistical hypothesis testing, and he generally opposed its use. Modern statistical hypothesis testing accommodates both types of hypotheses since the alternative hypothesis can be quite useful which will be pointed out later.

The basic issue is this: consider a large population with a known frequency distribution. Due to diminished knowledge of the population represented by a small subset of the population, a sample will generate a different frequency distribution than the parent population because the parameters estimated from the sample (a subset) will not

be the same as the parent population. That is the population and the subset can be characterized by the same frequency distribution but will have different distribution parameters. As the subset size increases, its frequency distribution more closely resembles that of the parent population and the estimated parameters approach the parent population parameters asymptotically. The likelihood ratio test and its close relationship to the χ^2 test can be used to determine what sample size will provide a reasonable approximation of the true population parameters. With this as background the relationship between parameter estimates, hypothesis testing, significance levels, the likelihood ratio test and the χ^2-distribution are outlined.

3.1 Hypothesis Testing

The two most common statistical inference procedures are confidence interval estimation and hypothesis testing. Confidence interval estimates provide a range of possible values within which the true, unknown population parameters reside. The values establishing the range for each distribution parameter are made and found acceptable from some quantifiable procedure. In contrast, when we make a declaration about the population parameters we are stating a hypothesis. For example, one can hypothesize that a true population parameter is equal to a specified value and do so without making a commitment a priori on whether the assumption is true or valid. The acceptance of the hypothesis is made based on data and how it relates to a relevant test statistic. The act of making inferences is known as hypothesis testing. A random sample (the data) is taken from the parent population and a statistical hypothesis is stipulated and tested. If the observations do not tend to support the hypothesis made, then the hypothesis is rejected.

On the other hand if the observations support the hypothesis, the hypothesis is not rejected, this does not necessarily mean that the hypothesis is accepted. It must be pointed out that hypotheses are not made in regards to estimated parameters, but rather about the true distribution parameters of the parent population. However, the estimated parameters are utilized as part of the process of drawing inferences.

Making a decision regarding a hypothesis can be considered a statistical event with an associated probability. Linked with the decision to accept or reject a hypothesis is a concept referred to as the level of significance, which is represented by the parameter α. Hypothesis testing is a procedure for rejecting a statement or not rejecting it, and one needs an ability to assess the probability of making incorrect decisions of either kind. Fisher (1925a) established a method to quantify the amount of evidence required to accept that an event is unlikely to occur by chance and defined this as the significance level. Significance levels are different then confidence levels but the two are related. Designating the confidence level as γ, the significance level and the confidence level are functionally related through the following simple expression

$$\gamma = 1 - \alpha \tag{3.1}$$

The confidence level when it is applied to parameter estimation is associated with a range, or more specifically with bounds or an interval, within which a true population parameter resides. Thus the confidence level, and by the equation above the significance level, is chosen arbitrarily based on the design application at hand and they impact the economics of the component design. The "significance level" is defined as the probability of mistakenly rejecting a hypothesis when the hypothesis is valid.

Establishing a value for either of this parameter effectively removes subjectivity from the decision making.

In the previous chapter the basic idea of estimating distribution parameters from strength data was presented. The quantities computed in that chapter are point estimates. The assumption is made throughout this work that the population distribution can be characterized by the Weibull distribution. Two distribution parameters, identified here generically as the pair of population parameter (θ_1, θ_2), will be estimated from a sample using maximum likelihood estimators. There are no preconceived notions concerning the values of these parameters when performing the point estimation calculations. One is simply trying to ascertain the best possible values of these parameters from a sample containing limited information. In contrast, when testing hypotheses made regarding (θ_1, θ_2) there are preconceived notions concerning their values. These preconceptions are embodied in the "null hypothesis" and the "alternative hypothesis". The latter, designated as H_1 and the former as H_0. The null hypothesis is the hypothesis that is tested. The null hypothesis assumes that samples obtained from the parent population are truly random. If the null hypothesis is true then any observed difference between the values obtained from the sampling procedure and the value assumed for the null hypothesis is merely a consequence of sampling variation, i.e., the differences are random errors and not systematic errors.

As part of the hypothesis testing process an alternative hypothesis is declared. Formally we test the null hypothesis, but the alternative hypothesis turns out to be just as important in determining the confidence bounds. In order to test the "null and alternative hypotheses" we need a test statistic pertinent to the population distribution and the

distribution parameters. The random variables in this process are the point estimates and the test statistic. A rejection region is defined, and this is known as the critical region for the test statistic. The observed tensile strength data is used to determine if the computed value of the test statistic (not the parameter estimates) lies within or outside the rejection region and thus defines the size of the critical region. If the test statistic is within the rejection region then we say the hypothesis is rejected at the 100α percent significance level. If α is quite small, then the probability of rejecting the null hypothesis when it is true is quite small.

Consider for example that if for a true distribution parameter θ there is a need to test the null hypothesis that $\theta = 0$ against the alternative that $\theta \neq 0$, then this test can be performed by determining whether the confidence interval for θ contains the value of zero. In a more general fashion, we can test the hypothesis that a true distribution parameter is equal to some stipulated value other than zero, i.e., $\theta = \theta_0$ (in the context here θ_0 is a stipulated value) against the alternative that $\theta \neq \theta_0$. Under these hypotheses a confidence interval can be constructed that contains the true population parameters with a probability of $\gamma = (1 - \alpha)$ and this interval also contains the value θ_0. The null hypothesis that the true distribution parameter equals θ_0 is not rejected at a significance level α. As noted earlier we can make mistakes in rejecting the null hypothesis. In hypothesis testing procedure, two types of errors are possible.

1. *Type-I Error*: The rejection of the null hypothesis (H_0), when it is true. The probability of committing a *Type-I error* is denoted as α.

2. *Type-II Error*: The failure in rejecting hypothesis (H_0), when the alternative hypothesis (H_1) is true. The probability of committing a *Type-II error* is denoted as β.

For either situation we are wrong in our judgment of the null hypothesis, H_0.

Now consider the two situations where correct decisions have been made. In the first case the null hypothesis is not rejected and the null hypothesis is true. The probability of making this choice is $\gamma = (1 - \alpha)$. This is the same probability associated with the confidence interval for the true population distribution parameters discussed above. The second situation, the probability of correctly rejecting the null hypothesis, is the complement of a *Type II error*, i.e., $(1 - \beta)$. In statistics this is known as the power of the test of the null hypothesis. This important criterion is used to determine the sample size. So in the process of making a correct decision we require high values of $(1 - \alpha)$ and $(1 - \beta)$.

Since a decrease in one type of error increases the other type of error, we must design our decision rules in order to minimize both types of errors in an optimal fashion. The probability of a *Type-I error* is controlled by making α a suitable number, say 1 in 10, or 1 in 20, or something smaller depending on the consequence of making a *Type-I error*. Minimizing the probability of making a *Type-II error* is not straight forward. Here β is dependent on the alternative hypothesis, on the sample size n, and in addition, on the true value of the distribution parameters tested. As we see in the next section the alternative hypothesis is greatly influenced by the test statistic chosen to help quantify our decision.

Hence when hypothesis testing is applied to the distribution parameters (θ_1, θ_2), a statement of equality is made in the null hypothesis H_0. Achieving statistical significance is equivalent to accepting that the observed results (the point estimates) are plausible if the null hypothesis is not rejected. The alternative hypothesis does not in general specify any specific values for the true population parameters but helps us to establish bounds. This is important when a confidence ring is formulated for a distribution pair. The size of the ring can be enlarged or reduced based on two parameters under our control, the significance level α, and the sample size, N.

3.2 The Test Statistic: The Likelihood Ratio Test

The test statistic utilized in this work is a ratio of the natural log of two likelihood functions. In simple terms one likelihood function is associated with a null hypothesis and the other is associated with an alternative hypothesis. Using the general approach to hypothesis testing the null and alternative hypotheses are defined respectively as

$$H_0 : \theta = \theta_0 \tag{3.2}$$

$$H_1 : \theta = \theta_1 \tag{3.3}$$

As they are used here, the hypotheses describe two complementary notions regarding the distribution parameters and these notions compete with one another. In this sense the hypotheses can be better described mathematically as

$$H_0 : \theta \in \Theta_0 = (\theta_1, \theta_2, ..., \theta_r) \tag{3.4}$$

$$H_1 : \theta \in \Theta_0^c = (\theta_1, \theta_2, ..., \theta_r)^c \tag{3.5}$$

Where r corresponds to the number of parameters in probability density function. The notion is that θ_0 and θ_1 are scalar values whereas Θ_0 is a vector of distribution parameters. The likelihood function associated with each hypothesis is

$$L_0 = \prod_{i=1}^{n} f\ (x_i \mid \theta \in \Theta_0)$$

(3.6)

for the null hypothesis and

$$L_1 = \prod_{i=1}^{n} f\ (x_i \mid \theta \in \Theta_0^c)$$

(3.7)

for the alternative function. The likelihood function, L_0, associated with the null hypothesis is evaluated using the maximum likelihood parameter estimates. Since the sample population is assumed to be a two parameter Weibull distribution, then a vector of distribution parameters whose components are the MLE parameter estimates is identified as

$$(\theta_1, \theta_2) = (\tilde{\theta}_1, \tilde{\theta}_2)$$

(3.8)

Now

$$H_0 : \theta \in \Theta_0 = (\tilde{\theta}_1, \tilde{\theta}_2)$$

(3.9)

that is Θ_0 contains all the MLE parameter estimates, and

$$H_1 : \theta \in \Theta_0^c$$

(3.10)

with Θ_0^C representing all point estimates that are not MLE parameter estimates. In essence we are testing the null hypothesis that the true distribution parameters are equal to the MLE parameter estimates with an alternative hypothesis that the true distribution parameters are not equal to the MLE parameter estimates.

Now the likelihood functions are expressed as

$$\hat{L}_0 = \prod_{i=1}^{n} f\,(x_i \mid \tilde{\theta}_1, \tilde{\theta}_2)$$

<div align="right">(3.11)</div>

and

$$\hat{L}_1 = \prod_{i=1}^{n} f\,(x_i \mid \theta_1, \theta_2)$$

<div align="right">(3.12)</div>

A test statistic is introduced that is defined as the natural log of the ratio of the likelihood

functions, i.e.,

$$T = -2\ln\left(\frac{\hat{L}_1}{\hat{L}_0}\right)$$

<div align="right">(3.13)</div>

The Neyman-Pearson lemma states that this likelihood ratio test is the most powerful test

statistic available for testing the null hypothesis. We can rewrite this last expression as

$$T = 2\,(\hat{\mathcal{L}} - \mathcal{L})$$

<div align="right">(3.14)</div>

Where

$$\hat{\mathcal{L}} = \ln\left(\hat{L}_0\right) = \ln\left\{\prod_{i=1}^{n} f\,(x_i \mid \tilde{\theta}_1, \tilde{\theta}_2)\right\}$$

<div align="right">(3.15)</div>

And

$$\mathcal{L} = \ln\left(\hat{L}_1\right) = \ln\left\{\prod_{i=1}^{n} f\,(x_i \mid \theta_1, \theta_2)\right\}$$

<div align="right">(3.16)</div>

The natural log of the likelihood ratio of a null hypothesis to an alternative

hypothesis is our test statistic and its distribution can be explicitly determined. The test

statistic is then used to form decision regions where the null hypothesis can be accepted

or rejected. A convenient result due to Wilks (1938) indicates that as the sample size n

approaches infinity, the test statistic $-2ln\ (T)$ for a nested composite hypothesis will be

asymptotically χ^2-distributed. If one hypothesis can be derived as a limiting sequence of another, we say that the two hypotheses are nested. In our case the sample $X_1, X_2, \ldots X_n$ are drawn from a Weibull distribution under H_0. These same samples are used in the alternative hypothesis, H_1, and since their parent distribution is assumed to be a Weibull distribution under both hypotheses, then the two hypotheses are nested and the conditions are satisfied for the application of Wilk's theorem.

The test statistic is designed in such a way that the probability of a *Type I error* does not exceed a constant α, a value that we control. Thus the probability of a *Type I error* is fixed and we search for the test statistic that maximizes *(1 − β)* where again β is the probability of a *Type II error*. In our case where inferences are being made on parameters from a two parameter Weibull distribution the degree of freedom for the χ^2-distribution is one and the values of the χ^2-distribution are easily calculated. One can compute the likelihood ratio T and compare *−2ln (T)* to the χ^2 value corresponding to a desired significance level and define a rejection region. This is outlined in the next section. The ratio L_0 / L_1 of the two likelihood functions defined above should be low in the optimal critical region − a result of minimizing α and maximizing *(1 − β)*. The ratio is high in the complementary region. A high ratio corresponds to a high probability of a correct decision under H_0.

The likelihood ratio depends on parameter estimates and is therefore a function of the sample strength data $X_1, X_2, \ldots X_n$, and thus the likelihood ratio is random variable. The likelihood-ratio test infers that the null hypothesis should be rejected if the value of the ratio is too small. How small is too small depends on the significance level of the test, i.e., on what probability of *Type I error* is considered tolerable.

The numerator corresponds to the maximum probability of an observed outcome under the null hypothesis. The denominator corresponds to the maximum probability of an observed outcome varying parameters over the whole parameter space. The numerator of this ratio should be less than the denominator or we have chosen our estimation function badly. The likelihood ratio hence is between 0 and 1. Lower values of the likelihood ratio mean that the observed result was much less likely to occur under the null hypothesis as compared to the alternate. Higher values of the likelihood ratio mean that the observed outcome was more than or equally likely or nearly as likely to occur under the null hypothesis as compared to the alternate and the null hypothesis cannot be rejected.
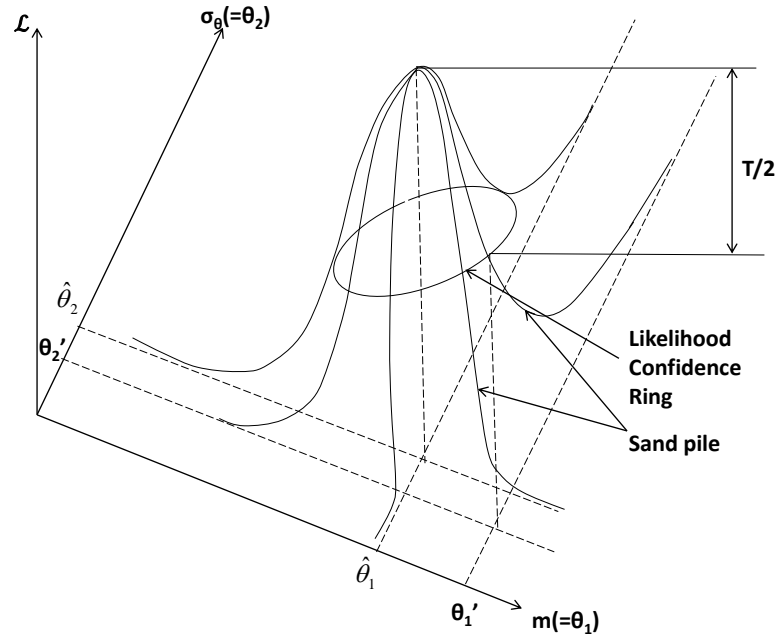

Figure 3.1 Likelihood Frequency Plot $\mathcal{L}(\theta_1, \theta_2)$ with Likelihood Confidence Ring and Associated Test Statistic T for $\theta_1 = \theta'_1$ and $\theta_2 = \theta'_2$

3.3 Likelihood Ratio Ring (Region of Acceptance)

The likelihood ratio confidence bound is based on the equation,

$$-2 \cdot \ln \left(\frac{L(\theta)}{L(\hat{\theta})} \right) \geq \chi^2_{\alpha;1} \tag{3.17}$$

Equation (3.17) is also written as

$$L(\theta_1, \theta_2) = L(\hat{\theta}_1, \hat{\theta}_2) \cdot \exp\left(\frac{\chi^2_{\alpha;1}}{2} \right) \tag{3.18}$$

The only unknowns in this equation are the generic parameters $\tilde{\theta}_1$ and $\tilde{\theta}_2$. The generic distribution parameters $\tilde{\theta}_1$ and $\tilde{\theta}_2$ are computed to satisfy equation (3.18). There is no closed-form solution, so an iterative numerical approach is used here. For a given significance level the confidence bounds are calculated by finding the extreme values of the contour ring in a plane perpendicular to the log likelihood axis (see Figure 3.1). A change in the significance level results in a different ring.

The iterative procedure is best conducted from inside the ring. To do this we use the maximum likelihood parameter estimates which are typically inside the ring. For a two parameter Weibull distribution equation (3.18) is now expressed as

$$L(m, \sigma_\theta) - L(\hat{m}, \hat{\sigma}_\theta) \cdot \exp\left(\frac{\chi^2_{\alpha;1}}{2} \right) = 0 \tag{3.19}$$

and one of the MLE parameter estimates, say $\hat{m}$, is used for m in the above equation to start the iterations. A pair of values is found for σ_θ. This procedure is repeated for a range of m-values until we have enough data points to saturate the ring which is then plotted.

CHAPTER IV

MATERIAL QUALITY ACCEPTANCE PROCESS WITH EXAMPLES

The material acceptance approach outlined here depends on several things. First one must have the ability to compute the probability of failure of the component under design. This will be designated $(P_f)_{component}$ and throughout this chapter this will be quantified using a hazard rate format, i.e., expressed as a fraction with a numerator of one. The hazard rate is simply the reciprocal of the probability of failure a quantity usually expressed as a percentage.

For monolithic ceramic materials the component probability of failure is modeled using a weakest link (series) reliability model. The underlying strength is characterized by a two parameter Weibull distribution. Thus a component probability of failure curve can be generated in an $m - \sigma_0$ graph. Points along the curve represent parameter pairs equal to a specified probability of failure. This curve will be referred to as a "Component Performance Curve." We overlay this graph with point estimates of the Weibull distribution parameters obtained from tensile strength data that a typical material supplier might have. Point estimates from this data to the right of the "component design curve" represent lower probability of failures. Conversely points to the left of this curve are

associated with performance curves with higher probability of failures. Thus one component design curve defines two regions of the $m - \sigma_0$ space, an acceptable performance region and a rejection region relative to a specified component probability of failure. See Figure 4.1 for an example.

Through Weibull size scaling the material characteristic strength σ_0 can be converted to a characteristic strength associated with a test specimen. For this effort the four point bend bar is chosen, and the characteristic strength is designated $(\sigma_\theta)_{specimen}$ in generals of the bend bar. Thus the component performance curve ($m - \sigma_0$ graph) can be converted to an equivalent test specimen performance curve plotted in an $m - (\sigma_\theta)_{specimen}$ graph. The reader is directed to Figure 4.2 for an example of this type of curve. Once again parameter pairs to the right of the material performance curve have lower probability of failures, and parameter pairs to the left are associated with higher probability of failures. So a choice is made a priori on the type of test specimen used to collect tensile failure strength.

Finally, the test performance curve, a curve representing a unique combination of component and test specimen, is married to a likelihood confidence ring that was discussed earlier. This allows the component fabricator to decide if the material supplier is providing a high enough quality material based on the component design and the failure data from four point bend tests (or other test specimen geometries). Keep in mind that parameter estimates are estimates of the true distribution parameters of the population, values that are never known in real life applications. However, through the use of the likelihood confidence ring method outlined in the previous chapter we can define a region in some close proximity of the estimated point parameter pair, and with

some level of assurance knowing that the true distribution parameters are contained within that region. If that region falls to the right of the test specimen performance curve the fabricator can accept the material based on the estimates of the Weibull parameters with a known level of quality, i.e., the significance level. Not surprisingly we define that particular procedure as the quality acceptance criterion. The details of this approach are outlined in this chapter using a simple thin wall pressure vessel as the component and as mentioned above, a four point bend bar as the test specimen of choice. The approach can easily extend to other components and other test specimen geometries.

4.1 Component Performance Curve – Example

For the state of stress in a thin wall pressure vessel

$$\sigma_1 = p \, r \, / \, t \tag{4.1}$$

$$\sigma_2 = p \, r \, / \, 2 \, t \tag{4.2}$$

$$\sigma_3 = 0 \tag{4.3}$$

Here p is the applied internal pressure, t is the wall thickness and r is a radius. The component probability of failure for a thin wall pressure vessel is given by

$$
\begin{aligned}
P_f \;=\; 1 \;-\; & \exp\left\{-\int\left[\left(\frac{\sigma_1}{\sigma_0}\right)^m + \left(\frac{\sigma_2}{\sigma_0}\right)^m\right]dA\right\} \\
\;=\; 1 \;-\; & \exp\left\{-\int\left[\left(\frac{pr_{avg}}{t\sigma_0}\right)^m + \left(\frac{pr_{avg}}{2t\sigma_0}\right)^m\right]dA\right\} \\
\;=\; 1 \;-\; & \exp\left\{-\left(\frac{pr_{avg}}{t\sigma_0}\right)^m\left[1+\frac{1}{2^m}\right]A_{comp}\right\}
\end{aligned}
\tag{4.4}
$$

Here r_{avg} is the average radius to the middle of the pressure vessel wall. With

$$\sigma_{max} = pr_{avg}/t \tag{4.5}$$

33

then the expression above becomes

$$P_f \;=\; 1 \;-\; \exp\left\{-\left(\frac{\sigma_{max}}{\sigma_0}\right)^m \left[1+\frac{1}{2^m}\right] A_{comp}\right\} \tag{4.6}$$

In the example presented here the pressure vessel geometry is

$$\text{inner radius} = 100 \text{ mm} \tag{4.7}$$

$$\text{thickness} = 5 \text{ mm} \tag{4.8}$$

$$\text{length of the vessel} = 400 \text{ mm} \tag{4.9}$$

and A_{comp} in equation (4.6) is the combined inner and outer area of the thin wall pressure vessel, i.e., failure can take place on either surface. Thus

$$A_{comp} \;=\; A_{inner} \;+\; A_{outer}$$

$$=\; 628.32 \text{ mm}^2 \;+\; 659.73 \text{ mm}^2$$

$$=\; 1{,}288.05 \text{ mm}^2 \tag{4.10}$$

and we are looking for a ceramic material whose Weibull distribution parameters yield a component probability of failure of

$$(P_f)_{component} = 1/500{,}000 = 2 \text{ x } 10^{-6} \tag{4.11}$$

In the above all values are arbitrarily chosen to illustrate the approach. The Weibull distribution parameter pairs that satisfy equation 4.6 with the geometry above are depicted as a curved line in Figure 4.1 below. If the probability of failure is increased the curve shifts to the right. Correspondingly if the probability of failure decreases the curve shifts to the left.
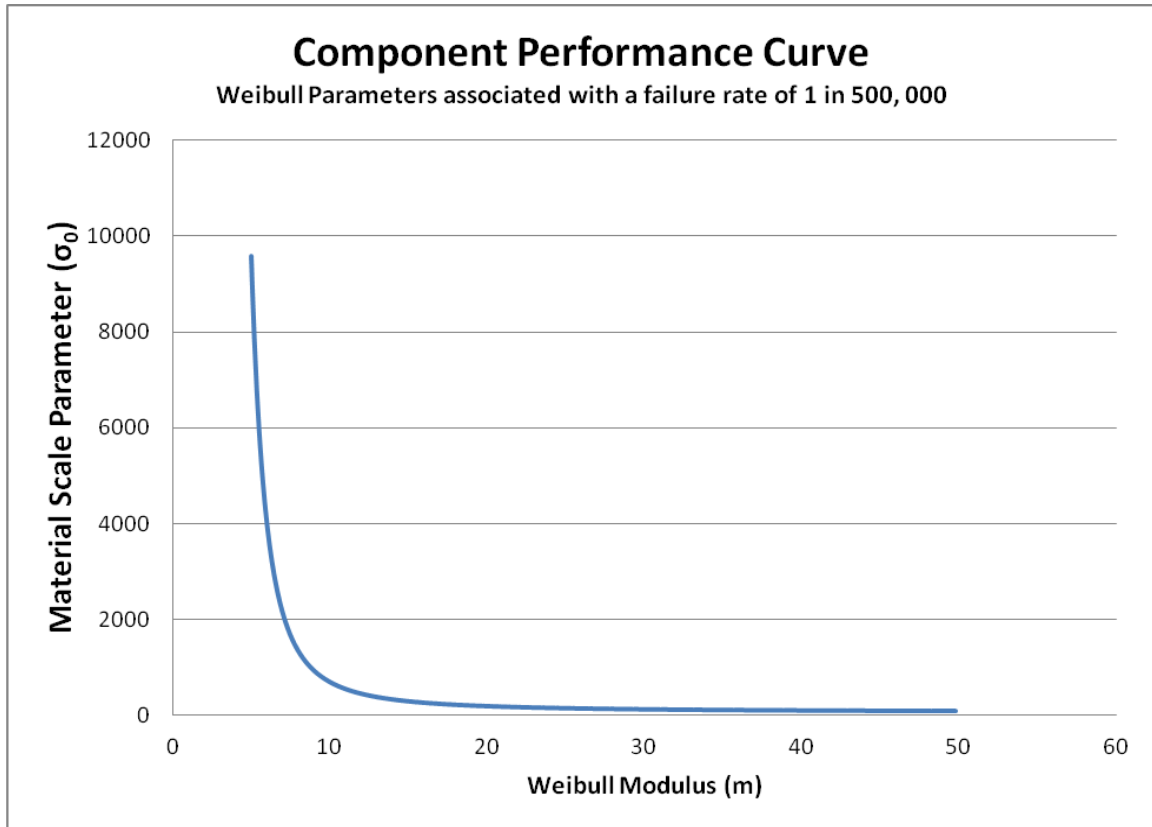
Figure 4.1 Component Performance Curve for a Thin Wall Pressure Vessel

4.2 Material Performance Curve

The next step in the quality assurance process is converting the information contained in Figure 4.1 to a test specimen configuration. Four point bend specimens tend to be the specimen geometry of choice to evaluate ceramic tensile strength due to alignment issues associated with tensile specimens. For this illustration a four point flexure specimen was selected in accordance with the guidelines set forth in ASTM C1161 configuration B. The ASTM standard for this specimen stipulates overall dimensions of 4 mm x 3 mm x 45 mm. The outer support span (L_o) of 40 mm (with 2.5 mm overhang on each side) and the inner support span (L_i) of 20 mm. Here we will identify the geometry of the cross section as

$$b = 4 \text{ mm} \tag{4.12}$$

and

$$d = 3 \text{ mm} \tag{4.13}$$

where b is the width of the specimen and d is depth of the specimen. The Weibull material characteristic strength parameter is related to the Weibull characteristic strength parameter through the expression

$$\sigma_0 = \left(A_{effective} \right)^{\frac{1}{m_A}} \left(\sigma_\theta \right)_A \tag{4.14}$$

Equation 4.14 is the expression that allows us to transform the component design information contained in Figure 4.1. For this application we assume that we will have a flaw population distributed along the tensile surface of the bend bar. If this is the case then the four point bend bar has an effective area of

$$A_{effective} = \left\{ L_0 \left[\frac{d}{(m_A + 1)} + b \right] \left[\left(\frac{L_i}{L_o} \right) m_A + 1 \right] \left[\frac{1}{(m_A + 1)} \right] \right\}^{1/m_A} \tag{4.15}$$

Inserting equation 4.15 into equation 4.14 yields

$$\sigma_0 = \left(\sigma_\theta \right)_A \left\{ L_0 \left[\frac{d}{(m_A + 1)} + b \right] \left[\left(\frac{L_i}{L_o} \right) m_A + 1 \right] \left[\frac{1}{(m_A + 1)} \right] \right\}^{1/m_A} \tag{4.16}$$

With equation 4.16 and the geometry of the pressure vessel geometry then equation (4.6) can be solved for σ_θ with assumed values of the Weibull modulus (m) and a specified failure rate P_f. These computed values of σ_θ are associated with a four point bend specimen. A plot of these parameter pairs is given in Figure 4.2 for the four point bend test specimen described above as well as the P_f stipulated in equation 4.11.
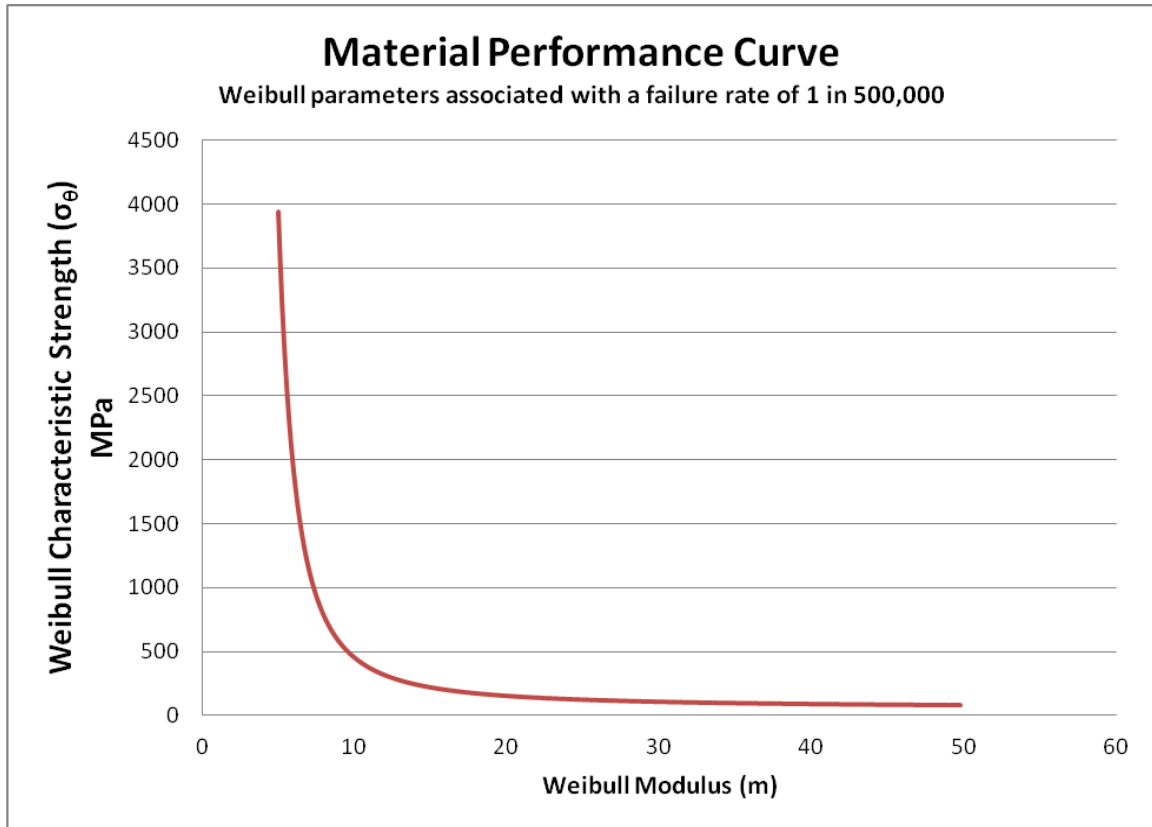
Figure 4.2 Material Performance Curve for Bend Bar Test Specimen

4.3 Likelihood Confidence Rings

In order to present aspects of the likelihood confidence rings we must utilize Monte Carlo simulation in order to obtain "test data." The details of Monte Carlo simulation as it applies to ceramic bend tensile data can be found in Palko (1992) and Hu (1994). Using Monte Carlo simulation allows us the knowledge of what the true distribution parameters are for a particular data set. These values can be arbitrarily chosen since the quality assurance process outlined here is invariant with respect to the Weibull distribution parameters. Here it is assumed that the Weibull modulus is 17 and the Weibull characteristic strength is 400 MPa. For a 90% confidence level and a sample size of 10, the likelihood confidence ring is plotted in Figure 4.3 along with the true

distribution parameters and the estimated distribution parameters. If the true distribution parameter pair was unknown we would be 90% confident that the true parameters are within the ring. If the Monte Carlo simulation process were continued nine more times, i.e., we were in possession of ten simulated data sets, then on average, one of those data sets would produce a likelihood confidence ring that did not contain the true distribution parameter pair.
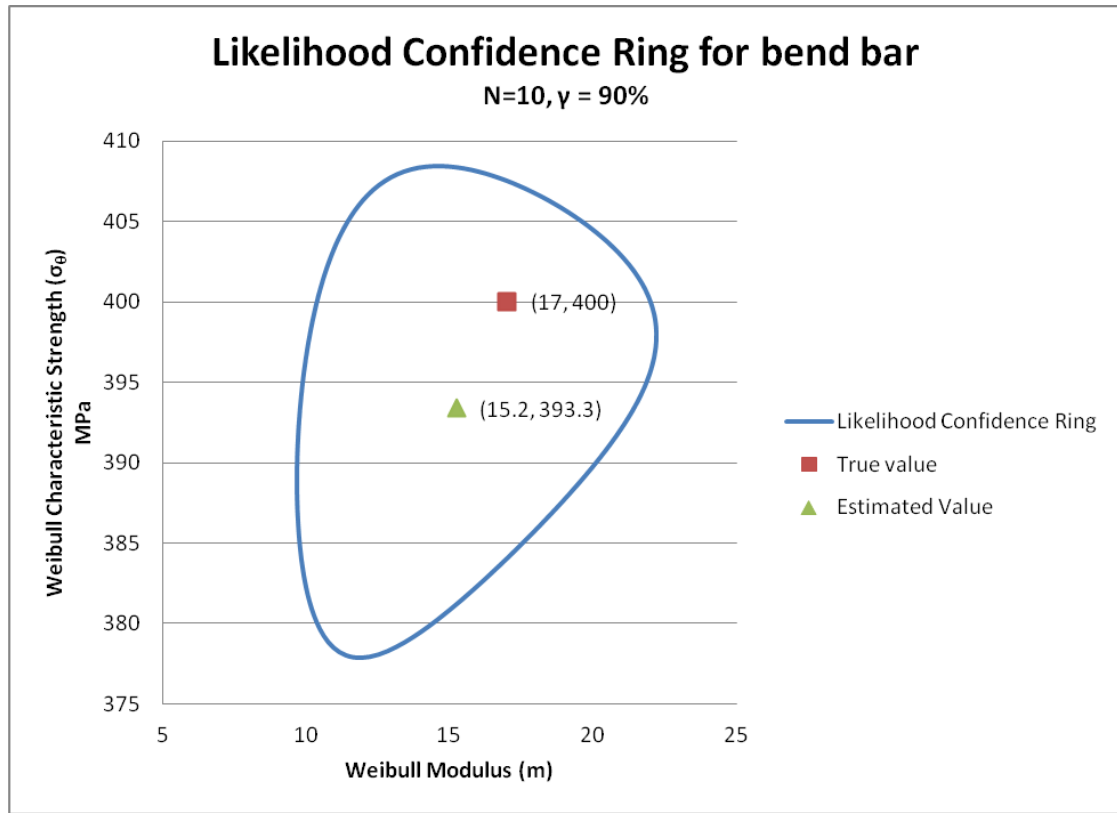

Figure 4.3 Confidence Ring Contour for a Sample Size of 10 and True Weibull Distribution Parameters ($m = 17$, $\sigma_\theta = 400$)

In the next figure the effect of holding the sample size fixed and varying the confidence level is presented. The result is a series of nested likelihood confidence rings. Here we have one data set and multiple rings associated with increments of the confidence level from 50% to 95%. Note that as the confidence level increases the size

of the likelihood confidence ring expands. For a given number of test specimens in a data set the area encompassed by the likelihood confidence ring would expand as we become more and more confident that the true distribution parameters are contained in the ring.
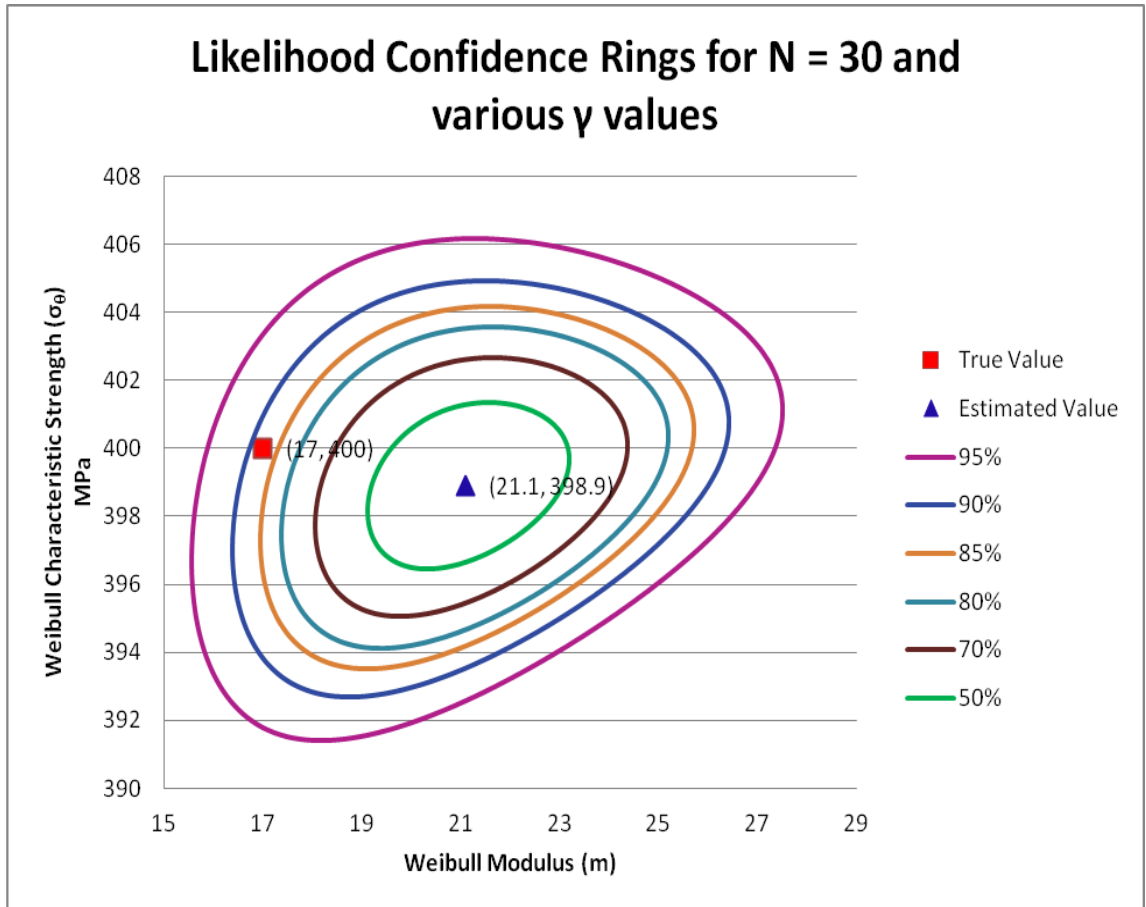

Figure 4.4 Likelihood Confidence Rings for 30 sample size with different γ values

The next figure, Figure 4.5, depicts the effect of varying the sample size and holding the confidence level fixed which is $\gamma = 90\%$. The sample size is increased from $N = 10$ to $N = 100$. Note that all the likelihood confidence rings encompass the true distribution parameters used to generate each sample. In addition the area within the rings grows smaller as the sample size increases. As the sample size increases we gain

information on the population and we thereby reduce the region that could contain the true distribution parameters for a given level of confidence.
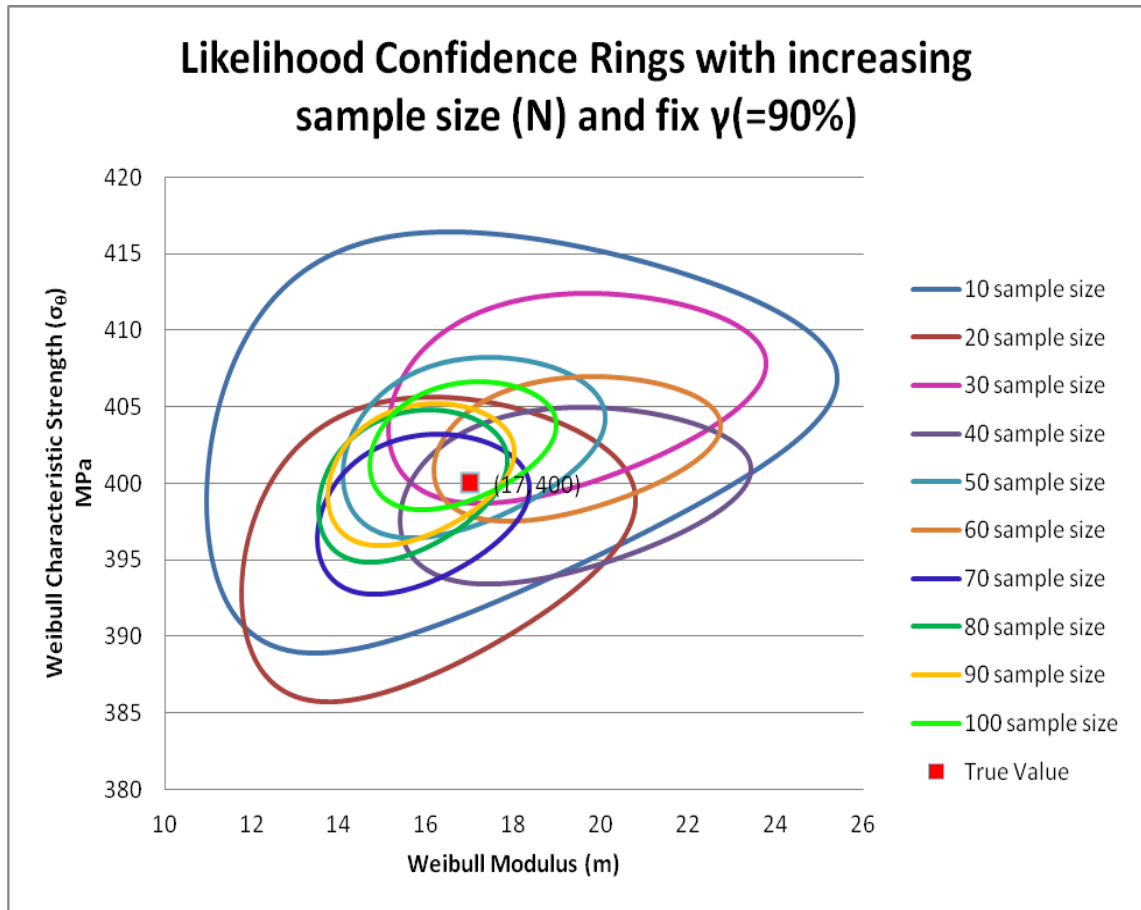

Figure 4.5 Likelihood Confidence Rings for 10 to 100 sample sizes and True Weibull Distribution Parameters ($m = 17$, $\sigma_\theta = 400$)

Figure 4.6 depicts a sampling procedure where the size of the sample is held fixed, i.e., $N = 10$, and the sampling process and ring generation has been repeated 100 times. For a fixed confidence level of 90% one would expect that ten rings would not encompass the true distribution parameters. Indeed that is the case. The 90 likelihood confidence rings that encompassed the true distribution parameters are outlined in blue. The 10 likelihood confidence rings that did not contain the distribution parameters are outlined in dark orange.
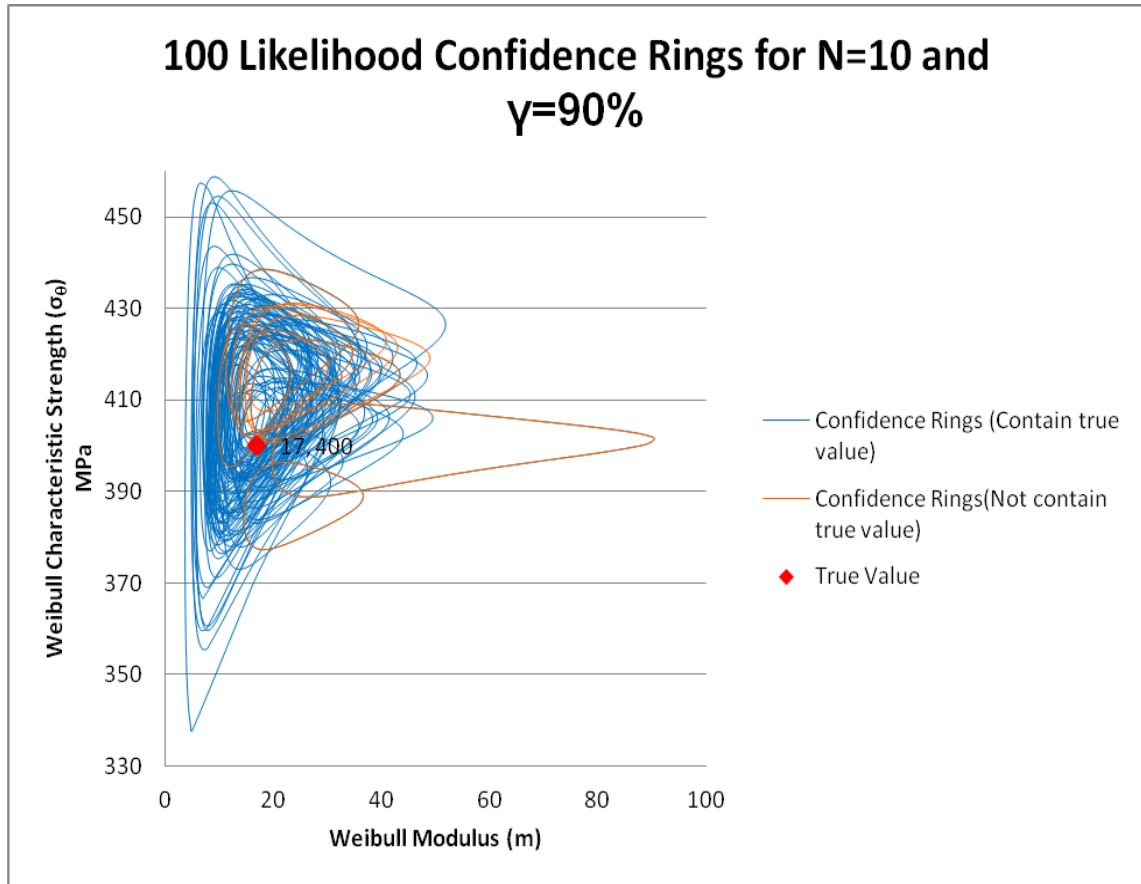
Figure 4.6 100 Numbers of Likelihood Confidence Rings for $N = 10$, $\gamma = 0.9$ and True Weibull Distribution Parameters ($m = 17$, $\sigma_\theta = 400$)

Finally we combine the two concepts, i.e., the likelihood confidence ring and the material performance curve. This is depicted in Figure 4.7. Here the material performance curve given in Figure 4.2 is overlain with the likelihood confidence ring from Figure 4.3. This is a graphical representation of the quality assurance process. Rings that reside completely to the right of the material performance curve would represent "acceptable" materials. Those rings to the left would represent unacceptable materials and would be rejected. In the specific case presented the material performance curve cuts through the likelihood confidence ring. In this case there are certain regions of the likelihood confidence ring that produce a "safe" design space and there is a region of

the likelihood confidence ring that produces an "unsafe" design space. In this situation we know the distribution parameters and they were purposely chosen to the right of the of the material performance curve. But given the sample size the ring did not reside entirely in the safe region. Moreover, in normal designs we never know the true distribution parameters so we do not know where the true distribution parameter pair resides inside the likelihood confidence ring.
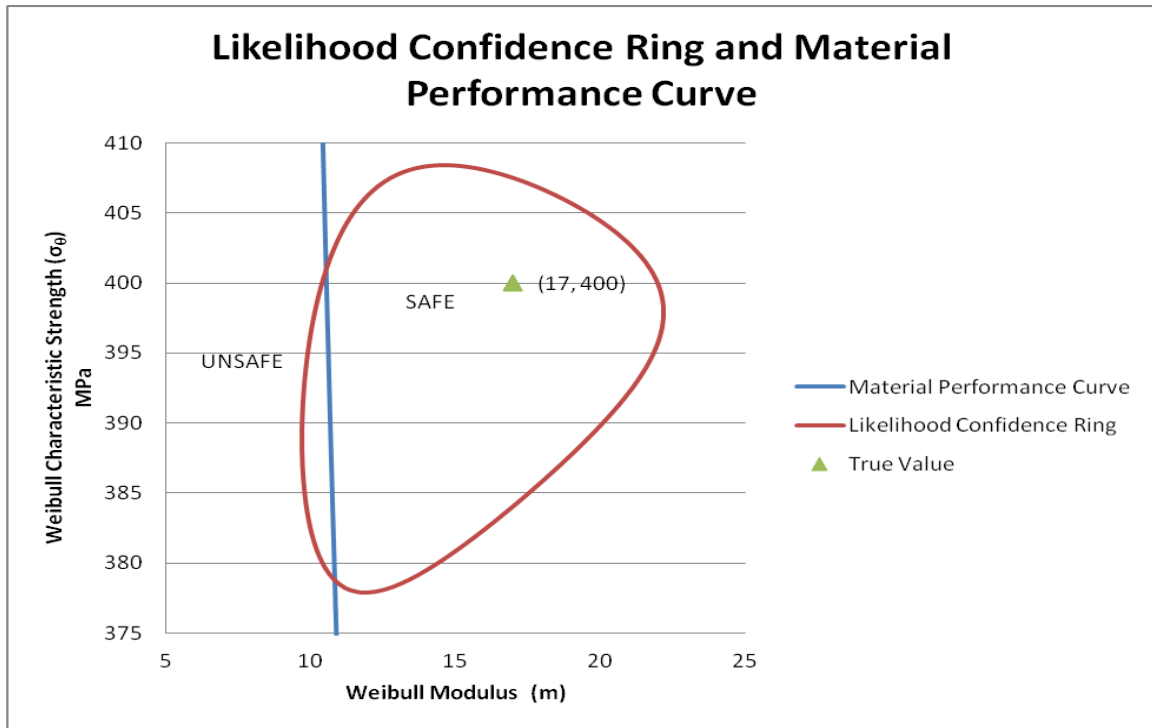

Figure 4.7 Material Performance Curve and Likelihood Confidence Ring Contour for *N*=10 and True Weibull Distribution Parameters (*m* = 17, *σ*$_\theta$ = 400)

For the case where the likelihood confidence ring resides totally to the left of the performance curve the choice to reject the material is quite clear. When the likelihood confidence ring lies completely to the right of the material performance curve, then once again, the choice is quite clear – accept the material. When the material performance curve slices through the likelihood confidence ring we can shift the material performance curve to the left, as is depicted in Figure 4.8. This shift represents a reduction of

component reliability, or alternatively an increase in the component probability of failure. Alternatively, the confidence bound associated with likelihood confidence ring can be reduced so the ring shrinks enough such that the ring is completely to the right of the material performance curve. This is depicted in Figure 4.9.
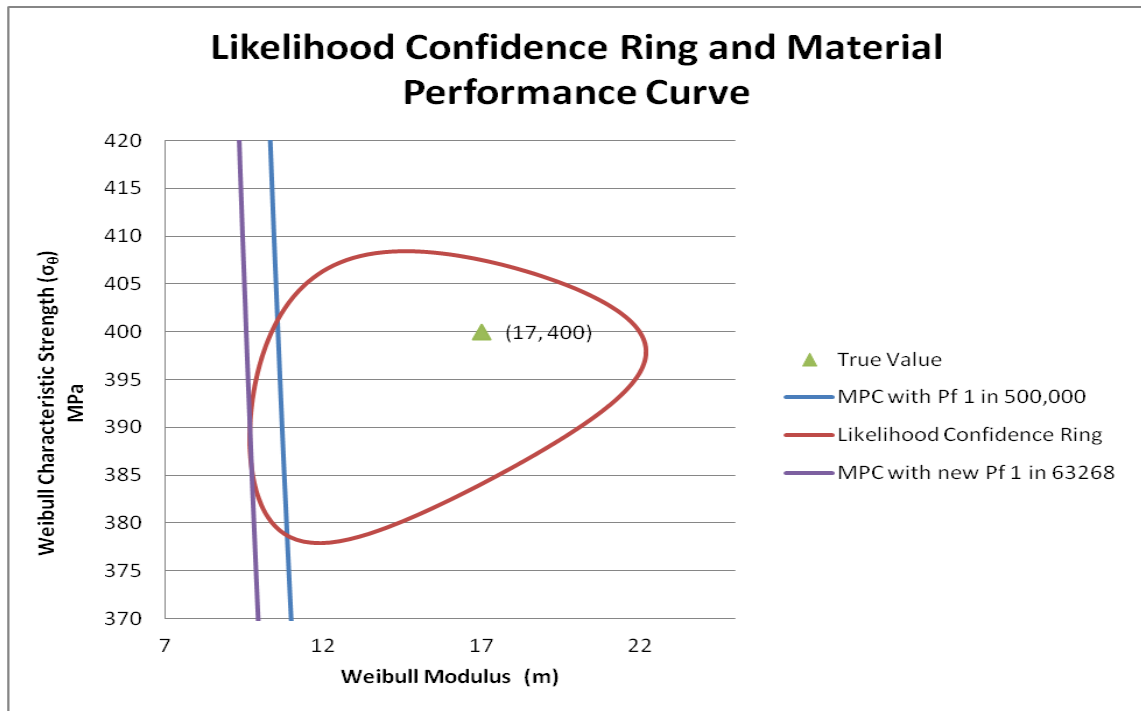

Figure 4.8 Two parallel Material Performance Curves and Likelihood Confidence Ring for True Weibull Distribution Parameters ($m = 17$, $\sigma_\theta = 400$)
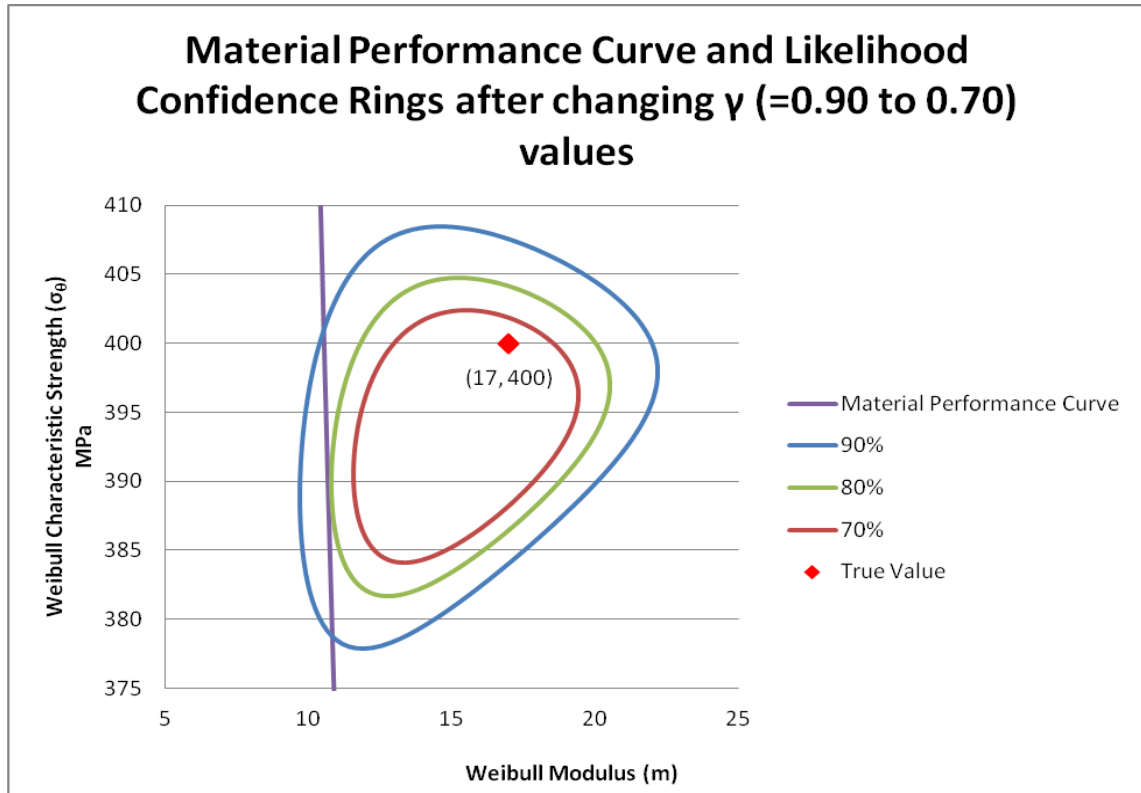
Figure 4.9 Material Performance Curves and Likelihood Confidence Rings after changing *γ* Values for True Weibull Distribution Parameters (*m* = 17, *σ*$_θ$ = 400)

An interesting aspect of this approach to quality assurance is that it seems that the likelihood confidence rings give a good indication of which side of the material performance curve the true distribution parameters lies. If the material performance curve slices through a likelihood confidence ring for a specified confidence level, then as the ring size is diminished the ring becomes tangent to one side of the curve or another. At the time this was written it was our experience that the side of the component reliability curve that the ring becomes tangent to matches with the side in which the true distribution parameters lie. It is emphasized that this is anecdotal. An example of where the true distribution parameters were chosen to the left of the material performance curve is depicted in Figure 4.10. The true distribution parameters are known because we are conducting a Monte Carlo simulation exercise to produce the data. As the confidence

level is reduced in Figure 4.10 the rings become tangent to the curve on the rejection side. This has not been rigorously proven, but it is an interesting observation from the interrogation of the process.
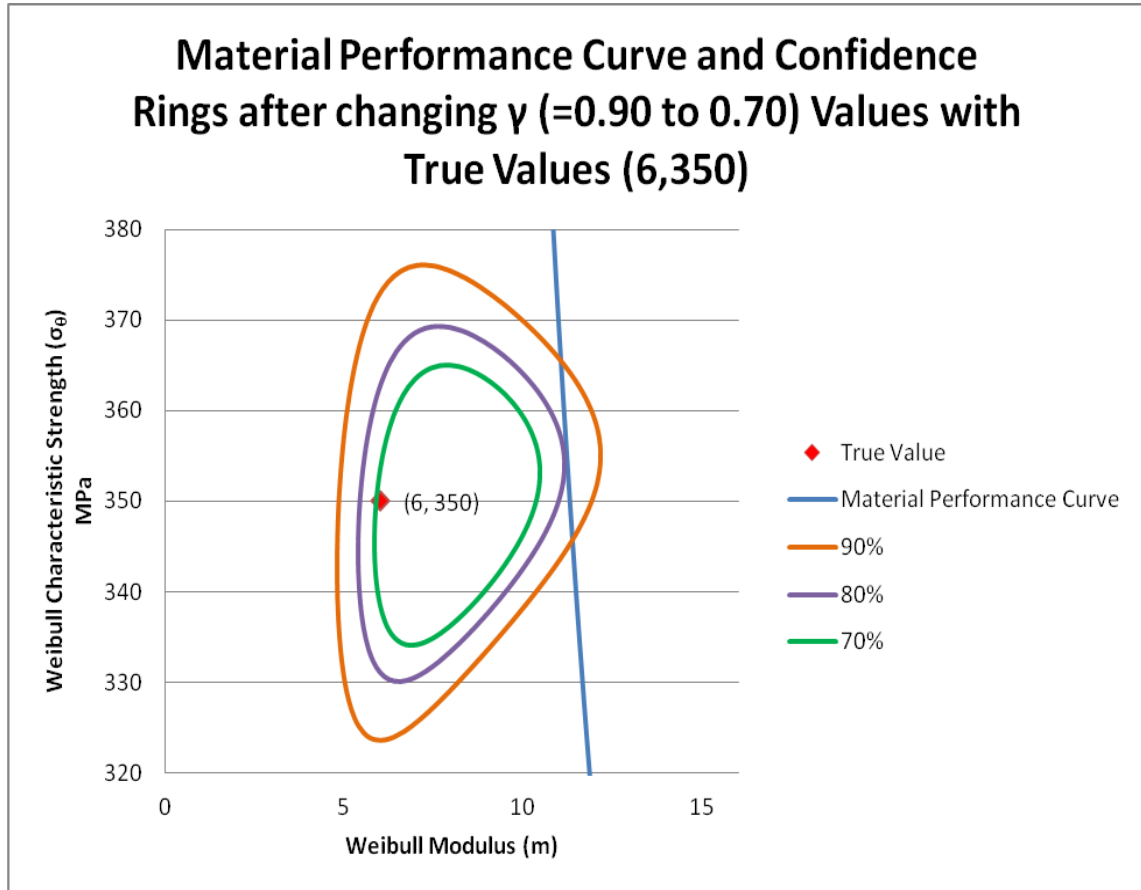

Figure 4.10 Material Performance Curves and Likelihood Confidence Rings after changing γ Values for True Weibull Distribution Parameters ($m = 6$, $\sigma_\theta = 350$)

CHAPTER V

SUMMARY AND FUTURE WORK

In summary, this effort focused on ceramic materials and the details associated with calculating point estimates for the Weibull distribution parameters associated with the tensile strength. One can easily generate point estimates from ceramic failure data using maximum likelihood estimators (recommended) or other estimators (e.g., linear regression or method of moments). What is not readily available is the ability to assess the quality of the point estimates. More information regarding the population, i.e., more data, will always improve point estimates; the question becomes how much data is sufficient given the application. The work outlined in this thesis speaks directly to this issue.

The details of calculating point estimates were presented and are summarized in Table 5.1. The concept of size scaling was introduced. The methodology is based on the well accepted practices outlined in ASTM C 1239. The system of equations needed to compute the Weibull parameters were presented in terms of a flaw population spatially distributed along the surface of the test specimen.

Hypothesis testing and the relationship it maintains with parameter estimation were outlined. A test statistic was defined which was then used to map out an acceptance region in the distribution parameter space. This provided the theoretical support for the equations used to generate the likelihood rings. Inferential statistics allowed us to generate confidence bounds on the true distribution parameters utilizing the test data at hand. These bounds are dependent on the size of the sample used to calculate point estimates. The effort focused on a particular type of confidence bound known as likelihood confidence rings.

Component reliability curves were discussed. A transformation method for size scaling available in ASTM C 1683 was utilized to demonstrate how a simple component, a thin wall pressure vessel, could have an attending component probability of failure curve presented in terms of the test specimen geometry. In the case presented here the test specimen was a four point bend bar. The concepts of the likelihood confidence rings and the component probability of failure curves in terms of a test specimen were combined graphically. This combination gives rise to a material qualification process. This process combines information regarding the reliability of the component as well as the parameter estimates to assess the quality of the material. A summary of the procedure is outlined in Table 5.1.

Table 5.1 Procedure For Establishing Performance Curves and Likelihood Confidence Ring

Steps	Procedure
1.	***Develop the "Component Performance Curve."*** Find the Weibull distribution parameter pairs (m, σ_0) that satisfy equation (4.6) for a specific component and stipulated probability of failure. In this thesis the structural component was a thin wall pressure vessel and the probability of failure was specified as $$P_f = 1{:}500{,}000$$ Plot the m - σ_0 curve (Figure 4.1) for a given component failure probability.
2.	***Establish the "Material Performance Curve."*** Convert the *Component Performance Curve* to specific test specimen geometry. In this thesis the four-point bend specimen was used and dimensions are chosen as per ASTM C1161 configuration B. Convert the material characteristic strength, σ_0, to a characteristic strength (σ_θ) associated with a given test specimen. Here the expression $$\sigma_0 = \left(A_{effective}\right)^{\frac{1}{m_A}}\left(\sigma_\theta\right)_A$$ was used for corresponding values of m. The effective area $A_{effective}$ was calculated from equation (4.15). Plot the m - σ_θ curve (Figure 4.2) for a given component failure probability.
4.	Estimate the Weibull parameters using maximum likelihood estimators. i.e., $$(\tilde{\theta}_1, \tilde{\theta}_2) = (\hat{m}, \hat{\sigma}_\theta)$$ from the strength data. Select a confidence level appropriate for the application $$\gamma = 1 - \alpha$$
5.	Use the test statistic T and compute pairs of (m, σ_θ) using the expression $$L\left(m, \sigma_\theta\right) - L\left(\hat{m}, \hat{\sigma}_\theta\right) \cdot \exp\left(\frac{\chi^2_{\alpha;1}}{2}\right) = 0.$$ Draw the likelihood confidence ring (Figure 4.3). Adjust the confidence ring such that it is tangent to the Material Performance Curve.

Looking beyond this effort, the approach outlined here would serve the nuclear graphite industry as well. Components fabricated from graphite will be used extensively in the next generation of nuclear reactors. Graphite is a brittle material in tension as well as in compression. However, the size effect exhibited by graphite materials is not characteristic of the size effect exhibited by ceramic materials. When a range of test specimen geometries are tested the mean strength from one size specimen to the next size does not faithfully follow the Weibull size scaling outlined in this work. There is a sense in the literature that density plays a role in impacting the size effect in graphite material. Mapping out an appropriate approach that would allow the methods outlined here to be adapted in qualifying components fabricated from graphite and utilized in nuclear reactors.

BIBLIOGRAPHY

1. ASTM Standard C1161-2002, "Standard Test Method for Flexural Strength of Advanced Ceramics at Ambient Temperature."

2. ASTM Standard C1239-2006, "Reporting Uniaxial Strength Data and Estimating Weibull Distribution Parameters for Advanced Ceramics."

3. ASTM Standard C1683-2008, "Size Scaling Of Tensile Strengths Using Weibull Statistics For Advanced Ceramics."

4. Fisher, R. A., "Theory of statistical estimation," Proc. Camb. Philos. Soc., 22, pg. 700-725, 1925a.

5. Fisher, R. A., "The mathematical distributions used in the common tests of significance," Econometrica 3, pg. 353-365, 1935.

6. Fisher, R. A. (1922b), "On the interpretation of χ^2 from contingency tables, and the calculation of *P. J. Roy. Statist. Soc.* 85," pg. 87-94; reprinted as Paper 19 in Fisher (1974).

7. Heger, A., Bruckner-Foit, A., and Munz, D., "STAU – A Computer Code to Calculate the Failure Probability of Multiaxially Loaded Ceramic Components," 2nd European Ceramic Society Conference, Euro-Ceramics, 1991.

8. Hu, J., "Modeling size effects and numerical methods in structural reliability analysis," MS thesis, Cleveland State University, Ohio, Nov 1994.

9. Nemeth, N.N., Manderscheid, J.M., and Gyekenyesi, J.P., "Ceramic Analysis and Reliability Evaluation of Structures (CARES) Users and Programmers Manual," NASA TP-2916, National Aeronautics and Space Administration, 1990.

10. Neyman, J. and Pearson, E., "On the Problem of the Most Efficient Tests of Statistical Hypotheses," Philosophical Transactions of the Royal Society of London, A231, pg. 289–337, 1933.

11. Palko, J. L., "An Interactive Reliability Model for Whisker Toughened Ceramics," Master's Thesis, Cleveland State University, 1992.

12. Pearson, K., "Contributions to the mathematical theory of evolution. II Skew variation in homogeneous material," Philosophical Transactions of the Royal Society of London, A186, pg. 343-414, 1895.

13. Pearson, K., "On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," Philos. Magazine, 5th Series 50, pg. 157–175, 1900a.

14. Wereszczak, A.A., Harper, D.C., Duty, C.E., Patel, P. "Glass Strengthening Via High-Intensity Plasma-Arc Heating," Journal of the American Ceramic Society, Vol. 93, Issue 5, pg. 1256-1259, May 2010.

15. Wilks, S.S., "The large sample distribution of the likelihood ratio for testing composite hypotheses," Ann. Math. Statist. 10, pg. 225-235, 1938.