Quantifying
Non-Destructive Evaluation
Capability and Reliability

Charles Annis
(407) 796-6565
Inspection Basics
NON DESTRUCTIVE EVALUATION

PROBABILITY OF DETECTION

CRACK SIZE

100%

0%
NDE Systems Classification

- "hit/miss" systems, which produce only qualitative information as to the presence or absence of a flaw.

- "a-hat vs. a" systems, which also provide some quantitative measure of the size of the indicated flaw.
Figure 1
Resolution in POD vs. Resolution in Cracksize

POD = n/N

Crack Found

Crack Missed

POD

Cracksize, a

0.00 0.05 0.10 0.15
MODELING POD FOR HIT/MISS DATA

\[ L(\theta; a, x) = \left[ \prod_{i=1}^{h} P_i \right] \left[ \prod_{j=1}^{n-h} (1 - P_j) \right] \]

where \( \theta = \begin{bmatrix} \mu \\ \sigma \end{bmatrix} \)

\( h = \) no. of hits

\( n = \) total no. of observations
The (log) Logistic Function

is sometimes used to model $POD(a)$.

$$POD(a) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

Where:

$x = \log(a)$, and

$(\beta_0, \beta_1)^T$ are model parameters.
The (log) Logistic Function

has interesting properties.

\[ g(x) = \ln\left( \frac{POD(a)}{1 - POD(a)} \right) \]

So that ...

\[ g(x) = \beta_0 + \beta_1 x \]

Nevertheless, the Normal Hazard function, used to model "a-hat vs. a" data, is also used to model "hit/miss" data, to facilitate comparisons between the different inspection techniques.
The $POD(a)$ Function

Relates the Probability of Detection to Cracksize, $a$

\[ POD(a) = 1 - Q \left( \frac{x - \mu}{\sigma} \right) \]

Where:

$Q$ is the Normal Survivor Function,
$x = \log(a)$, and
$(\mu, \sigma)^T$ are model parameters.
APPARENT CRACK SIZE IS RELATED TO ACTUAL SIZE

Legend
- △ Misses
- ■ Data
- □ Saturations
- ○ CENSORED FIT
\( \hat{a} \text{ vs. } a \) CAN BE TRANSLATED TO POD(a)
Elementary Statistical Concepts for NDE
Statistics for NDE Engineers

(Secrets that Statisticians don't want Engineers to know)
Statistics Overview (continued)

A "statistic" is a function of the data alone.

eg: \[ \bar{x} = f(x_1, x_2, \ldots, x_n) \]

\[ \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \]
Statistics Overview (continued)

Engineers are familiar with probability distributions of physical characteristics.
Secret #1

Statisticians work with probability distributions of population parameters.

(Actually, "estimates" of population parameters, i.e., statistics.)
A "Population Parameter" ...

- describes the underlying behavior of the data.
  eg: \((\mu, \sigma^2\) for Normal Distribution)

- is fixed (it has no distribution).
A "Population Parameter" ...

- is not a function of the observations, but can be estimated from them.

- these "Parameter Estimates" are functions of the data. (more on this later)
"Probability" is a statement about frequency of occurrence.
Here's the Plan:

We'll discuss a simple univariate example, then generalize to the multivariate case used to quantify NDE performance.
A Review of Probability

for people who hate Statistics:

\[ P(x) = \frac{n_x}{N} \]

where: \( n_x \) is number of "successes,"
\( N \) is total number of opportunities,
and a "success" has trait \( x \).
Probability Review, cont.

$P(x)$ is the relative frequency of occurrence of trait $x$.
(eg: $X = $ fatigue life of $x$ cycles)

\[
p = \frac{10}{90+10} = 10\%
\]

VENN diagram
Probability Review, cont.

A "Probability Distribution" describes the relationship between $x$ and the relative frequency of $x$, $f(x)$.
Probability Review, cont.

example:

If $X \sim N(\mu, \sigma^2)$

Then,

\[
P(\mu - 1\sigma < x < \mu + 1\sigma) \approx 63\%
\]

\[
P(\mu - 2\sigma < x < \mu + 2\sigma) \approx 95\%
\]

\[
P(\mu - 3\sigma < x < \mu + 3\sigma) \approx 99\%
\]
Probability Review, Summary

Population parameters (eg: $\mu, \sigma^2$) tell how often to expect a given value of $x$.

$$\text{Probability} = P(x \mid \theta)$$

eg: $\theta = \begin{bmatrix} \mu \\ \sigma \end{bmatrix}$
Secret #2

The "Likelihood Function" permits "estimating" Population Parameters from the behavior of the data.

$$\text{Likelihood} = L(\theta \mid x)$$
Likelihood describes the behavior (likelihood) of the population parameter estimates, given the data.

ie: the parameter estimates are a function of the observations
Likelihood
describes the behavior (likelihood) of the population parameter estimates, given the data.

ie: the parameter estimates are a function of the observations
Likelihood has a maximum
(for well-behaved situations)
MLEs (Maximum Likelihood Estimators) are Normally Distributed.

(see example)
MLE Example:

- The sample mean, $\bar{x}$, is a mle and has a Normal Distribution with mean $\mu$ and variance $\sigma^2/n$, where $n$ is the sample size.

- For large samples this is true even when the population distribution is NOT-Normal.
Secret #3, comment

It is the population parameter estimates (not the population values) who's normal behavior can be exploited (for example, to place "confidence-limits" on \( POD(a) \)).
MLEs are asymptotically Normal with mean $\mu^{[1]}$ and variance-covariance matrix determined $^{[2]}$ from the Likelihood function.

notes:

[1] $\mu = \text{population mean vector}$

[2] $V = \left[ \frac{\partial^2 L}{\partial \theta^2} \right]^{-1}$
Using Secret #3
So what?

We can estimate $POD(a)$ behavior from the experimental data to provide a mean $POD(a)$ relationship and construct confidence limits.
POD(a) can be described quantitatively

Legend
Mean Behavior
Confidence Limit