3

Application of Statistical Analysis in High Voltage Engineering

3.1 Introduction

High voltage engineering relies heavily on experimental data. However, experiments are often costly and time consuming, and in this context, analysis of experimental data has always been of great importance to extract proper information. Because statistical analysis is independent of the phenomenon or objects under investigation, it is applicable to a variety of subjects to sort out their representative parameters and to derive models that help make predictions about their performance under different conditions. It is no surprise therefore, that statistical analysis has found its way into high voltage engineering as a means to plan testing and analyze experimental data (Hauschild and Mosch, 1992).

However, precisely because it is phenomenon-independent, statistical analysis alone cannot provide a physical explanation on the phenomena under study. It can only present a better, more coherent picture of the set of available data, as well as provide some insight into the phenomenon investigated. Best results are obtained when statistical analysis is paired with an understanding of the physics involved.

Statistical applications to engineering studies are numerous. In the field of high voltage engineering, it has been used to

- Characterize data, for example, to estimate low-probability withstand levels.
- Verify an experimental law.
- Extrapolate the test results on small samples to full-scale insulation system by accounting for length, surface and volume effect, using weak-link statistics.
- Optimize experimental data to separate the component distributions in the data, to improve the amount of experimental data available.
- Estimate the risk of failure and assess breakdown voltage-time characteristics, using joint probability.
- Plan and elaborate more efficient test methodologies aimed at determining low-probability withstand levels.

3.2 Characterization of a Set of Data under Specific Conditions

Anyone who conducts experimental tests is familiar with the fact that repeating the same test many times does not give the same results. Generally the data will be scattered over a finite range of values. This is illustrated by the plot in Figure 3.1. It shows the variations of the ac breakdown field, $E_c$, of a bushing shield–plane gap in oil, in chronological order. It can be seen that though the breakdown field varies from one measurement to the next, it remains within a finite range, bounded by the values 3.1 and 6.5 kV/mm.
3.2.1 Global Statistical Parameters

Two very important parameters are used to describe a set of experimental data points. Those are the **mean value** $U_m$ and the **standard deviation** $\sigma$, defined according to (Miller and Freund, 1965)

\[
U_m = \frac{1}{N} \sum u_i
\]  

and

\[
\sigma^2 = \frac{1}{N-1} \sum (u_i - U_m)^2
\]

Here, $u_i$ are the experimental measurements and $N$ is the total number of measurements. The mean value $U_m$ is representative of the population of experimental data, while the variance $\sigma^2$, and related standard deviation $\sigma$, represents the degree of dispersion of the measurements with respect to the mean and is indicative of the reproducibility of the results. For the set of results in Figure 3.1, the mean breakdown field is $E_m = 4.43$ kV/mm, and the standard deviation is $\sigma = 0.74$ kV/mm, with the coefficient of variation $\sigma/E_m = 0.175$ or 17.5%.

3.2.2 Experimental (Discrete) Distribution

When the volume of experimental data is sufficient, statistical analysis provides an elegant means to evaluate these data, basically by deriving an experimental distribution relating the frequency of occurrence of the measured data expressed as a probability, $P(u)$. Two cases may be distinguished according to the total amount of data gathered.

3.2.2.1 Experimental Distribution with a Small Amount of Data

For most experimental studies, the total number of data is limited, typically in the order of a few tens of measurements. To derive an experimental distribution from a set of breakdown voltage measurements,
one first classifies the results in increasing order of the breakdown voltages. The cumulative probability that the breakdown voltage will be lower than or equal to a specific level \( U \) is then defined according to

\[
P(u \leq U) = \frac{n(u \leq U)}{N + 1}
\]

where
- \( n \) is the number of breakdowns occurring at voltages \( u \leq U \)
- \( N \) is the total number of measurements

The use of \((N + 1)\) in the preceding definition of the cumulative probability \( P(U) \) ensures that the experimental probability distribution is

- Continuously increasing
- Comprised between 0 and 1, more precisely between limiting values of the relative frequencies of \( 1/(N + 1) \) and \( N/(N + 1) \)
- The range of probabilities covered by the experimental data is symmetrical with respect to the median value \( U_{50\%} \)

Figure 3.2 shows the experimental distribution derived from the earlier data set. It can be seen that with a total number of 29 measured breakdown fields, the experimental cumulative distribution is defined between the extreme values of probabilities \( P_{\text{min}} \) and \( P_{\text{max}} \):

\[
P_{\text{min}} = \frac{1}{N + 1} \pm 3.333 \times 10^{-2} \quad \text{and} \quad P_{\text{max}} = \frac{N}{N + 1} \pm 9.667 \times 10^{-1}
\]

FIGURE 3.2 Experimental cumulative probability distribution of the ac breakdown field, \( E_c \), of a 20 cm bushing shield–plane gap in oil.
3.2.2.2 Experimental Distribution with a Large Amount of Data

Statistical analysis is also useful in cases where there is a large amount of data, for example, as is obtained with an automatic data acquisition system over long periods of time. One may be interested to limit the representation of the set of data to \( m \) data points, with \( m \leq N \), defined as follows:

- Define the extreme limits of the data \( U_{\text{min}} \) and \( U_{\text{max}} \), and reclassify the \( n \) experimental data in increasing order.
- Define \( m - 1 \) equal intervals, \( \Delta u = (U_{\text{max}} - U_{\text{min}})/(m - 1) \), with \( U_{\text{min}} \) and \( U_{\text{max}} \) becoming the first and last values.
- The probability density function is defined by \( m - 1 \) values \( p \):

\[
p_j = \frac{dn}{du} = \frac{n_j}{n + 1}
\]

where \( n_j \) is the number of measurements comprised in the \( j \)th interval.
- The cumulative probability function is defined by \( m \) values \( P_j \):

\[
P_j(u \leq u_j) = \frac{1}{n + 1} \sum_{k=1}^{j} n_k
\]

with \( j = 1, m \), and \( \sum n_k \) is the number of measurements comprised in the first \( j \) intervals by increasing order.

Figure 3.3 represents the cumulative distribution of corona losses (CL) derived from long-term recordings made at IREQ during the evaluation of the corona performances of bundle conductors considered for HVDC transmission lines operating at ±750 kV dc, using the earlier method (Maruvada et al., 1977).
3.2.2.3 Correspondence with the Global Parameters

With the experimental probability density and cumulative probability functions defined earlier, the following correspondence exists with the global parameters defined earlier:

The expectation (mean value)

\[ U_m = \sum p_j \mu_j \]  \hspace{1cm} (3.7)

The variance

\[ \sigma^2 = \sum (p_j \mu_j - U_m)^2 \] \hspace{1cm} (3.8)

with \( \sigma \) as the standard deviation.

3.2.2.4 Other Characteristic Parameters

The experimental distributions of the probability density and cumulative probability also allow the definition of several other characteristic parameters, which are currently used in practice, namely,

- **Most probable voltage, \( U_p \):** Is the voltage at which the probability density function is maximal.
- **Median voltage, \( U_{50\%} \):** Is the voltage at which the cumulative probability of occurrence is equal to 0.5.
- **Withstand voltage of \( x\% \):** Is the voltage at which the cumulative probability of occurrence is equal to \( x\% \).
- **Confidence interval at \( y\% \):** Is defined as the range of voltages \( u \), which ensures a probability of occurrence of \( y\% \). The confidence interval is limited by the withstand voltages of \((0.5y)\%\) and \((100-0.5y)\%\) and represents a measure of the reproducibility of the phenomenon investigated.
- **Coefficient of variation:** Is defined as the ratio \( \sigma/U_m \).

3.2.3 Statistical Functions (Models)

Because of the limitations of time and cost, experimental data are often obtained under laboratory conditions. As good as they are, laboratory conditions never fully correspond to the operating conditions of the equipment. Therefore, there is a great incentive to extrapolate the results beyond the conditions under which they were obtained. A useful approach consists in defining an analytical distribution function that models the experimental distribution expressed by Equation 3.3 or 3.6, which also satisfy the general conditions for a cumulative probability distribution:

- To be continuously increasing over the range of variations of the independent variable
- To have a value comprised between 0 and 1

By appropriate uses of interpolation or extrapolation, the statistical model thus obtained can be used to determine the behavior of the phenomenon under investigation, for ranges not covered by the experimental data. Several statistical functions exist that are currently used in practice. The most common are presented in detail in the succeeding texts. They are classified in two general categories of discrete and continuous distributions as related to the nature of the variable.

3.2.3.1 Discrete Distributions

In discrete distributions, the variable can take only discrete values, denoted by \( X_i \), for which the relative frequency is defined. Typical discrete distributions are uniform, binomial, and Poisson.
3.2.3.1.1 Uniform Distribution

The random variable can assume \( n \) different values:

\[
\text{Individual probability } \quad p_i = \frac{1}{n} \quad (3.9)
\]

\[
\text{Cumulative probability } \quad P(X \leq X_i) = \sum p_i \quad (3.10)
\]

\[
\text{Parameters } \quad n, X_i (i = 1, \ldots, n)
\]

\[
\text{Expectation } \quad EX = \frac{1}{n} \sum X_i \quad (3.11)
\]

\[
\text{Variance } \quad D^2 X = \frac{1}{n} \sum (X_i - EX)^2 \quad (3.12)
\]

3.2.3.1.2 Binomial Distribution

Binomial distribution applies to the case where the random variable is of Bernoulli type, which can take only one of the two values, 1 and 0, corresponding to the occurrence or nonoccurrence of an event \( E \), for example, breakdown of a test gap. Defining the random variable \( X \) as the number of times \( E \) occurs in \( n \) independent trials, with values 0, 1, \ldots, \( n \), one gets the following expressions:

\[
\text{Probability density function } \quad p(k) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad (3.13)
\]

\[
\text{Cumulative probability } \quad P(X \leq k) = \sum_{m=0}^{k} \binom{n}{k} p^m (1 - p)^{n-m} \quad (3.14)
\]

\[
\text{Parameters } \quad n \quad \text{and } p
\]

\[
\text{Expectation } \quad EX = np \quad (3.15)
\]

\[
\text{Variance } \quad D^2 X = np(1 - p) \quad (3.16)
\]

On the other hand, from a given set of experimental data, one can define the most representative statistical model from

Point estimate for \( p \): \( p = k/n \), the relative frequency from \( n \) trials in which event \( E \) occurs \( k \) times.

Confidence estimate for \( p \):

- The \( F \) distribution produces the one-sided confidence region with an upper limit as

\[
p_0 = \frac{(k + 1) F_{m1,m2,x}}{n - k + (k + 1) F_{m1,m2,x}}
\]

and the confidence interval

\[
I = [0, p_0]
\]

where \( F_{m1,m2,x} \) is the quintile of \( F \) distribution of order \( q = \varepsilon \), the confidence coefficient and with degrees of freedom \( m_1 = 2(k + 1) \) and \( m_2 = 2(n - k) \)
and one-sided confidence region with a lower limit as

\[ p_1 = \frac{k}{k + (n - k + 1) F_{m_3,m_4,x}} \]

and the confidence region

\[ I = [p_1, 1] \]

with \( m_3 = 2(n - k + 1); m_4 = 2k \).

- A two-sided confidence region after Neumann is tabulated in Muller et al. (1973) for \( n = 1 \) to \( n = 30 \). These confidence limits are represented in Figure 3.4. Larger values of \( n \) are covered by the asymptotic relation

\[ p_0 = \frac{1}{n + \lambda^2} \left( k + \frac{\lambda^2}{2} \mp \lambda q \sqrt{\frac{k(n - k)}{n} + \frac{\lambda^2}{4}} \right) \]

and the confidence interval

\[ I = [p_0, p_1] \]

where \( \lambda_q \) is the quantile of the standardized normal distribution of order \( q = (1 + \sigma)/2 \).

### 3.2.3.1.3 Poisson Distribution

Poisson distribution is derived from the binomial distribution when \( n \) tends towards infinity, and at the same time, \( n_p = \lambda \) remains constant:

\[ \lim_{n \to \infty, np = \lambda} \binom{n}{k} p^k (1 - p)^{n-k} = \frac{\lambda^k}{k!} \exp(-\lambda) \quad \text{with} \quad k = 1, 2, \ldots \]

Since the probability \( p \) is very small for large \( n \) and \( n_p = \lambda \) = constant, the Poisson distribution describes rare events.

**Probability density function**

\[ p(k) = P(x = k) = \frac{\lambda^k}{k!} \exp(-\lambda) \quad (3.17) \]

**Cumulative probability**

\[ P(X \leq k) = \sum_{m=0}^{k} \frac{\lambda^m}{m!} \exp(-\lambda) \quad (3.18) \]

**Parameter** \( \lambda \)

**Expectation**

\[ EX = \lambda \quad (3.19) \]

**Variance**

\[ D^2X = \lambda, \text{equal to the expectation } EX \quad (3.20) \]

**Point estimate for \( \lambda \)**

\[ \hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad (3.21) \]
Confidence estimate for $\lambda$

- One-sided confidence intervals
  \[ I = \left[ \frac{\lambda_{k,x}}{n} , \infty \right] \quad \text{and} \quad I = \left[ 0 , \frac{\bar{\lambda}_{k,x}}{n} \right] \]

- Two-sided confidence intervals
  \[ I = \left[ \frac{\lambda_{k,(1+e)/2}}{n} , \frac{\bar{\lambda}_{k,(1+e)/2}}{n} \right] \]

where $\lambda_{k,x}$ and $\bar{\lambda}_{k,x}$ are, respectively, the lower and upper confidence limits.

The confidence factors are related to the quantile of the $\chi^2$ distribution.

### 3.2.3.2 Continuous Distributions

In continuous distributions, the variable can take any values, the relative frequency of which is defined by a density function $f(x)$. Typical continuous distributions are Maxwell, normal, extreme values, and Weibull.
3.2.3.2.1 Maxwell Distribution

Maxwell distribution was introduced by J.C. Maxwell to describe the distribution of velocities of gas molecules under equilibrium conditions, presented in Chapter 4. Maxwell distribution is one of the rare cases where the statistical distribution is derived directly from a physics theory, in this case the kinetic theory of gases. However, it is rarely used as a statistical model, being too specifically identified to gaseous kinetic parameters. The probability density of the velocity of the gas molecules, \( f(v) \), can be expressed in terms of the mass and temperature of the gas according to

\[
f(v) = \frac{4}{\sqrt{\pi}} \left( \frac{M}{2kT} \right)^{3/2} v^2 \exp \left( -\frac{Mv^2}{2kT} \right) \quad (3.22)
\]

The cumulative Maxwell distribution is

\[
P(v) = \int_{-\infty}^{v} f(v) \, dv = \frac{4}{\sqrt{\pi}} \left( \frac{M}{2kT} \right)^{3/2} \int_{-\infty}^{v} v^2 \exp \left( -\frac{Mv^2}{2kT} \right) \, dv \quad (3.23)
\]

where

- \( k \) is the Boltzmann constant
- \( v \) is the velocity
- \( M \) is the mass of the molecule
- \( T \) is the absolute temperature

One can obtain the following characteristic parameters:

- Most probable velocity \( c_p = \sqrt{\frac{2kT}{M}} \)
- Average velocity \( c_v = \frac{2}{\sqrt{\pi}} c_p = 2 \sqrt{\frac{2kT}{\pi M}} \)
- Effective velocity \( c_{rms} = \sqrt{\frac{3}{2}} c_p = \sqrt{\frac{3kT}{M}} \)

3.2.3.2.2 Normal Distribution

Of the various distribution functions, the *normal* function is probably the most familiar. The probability density, \( f(u) \), and the cumulative probability, \( P(u) \), are defined according to (Miller and Freund, 1965)

\[
f(u) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{u - U_m}{\sigma} \right)^2 \right] \quad (3.24)
\]

\[
P(u) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{u} \exp \left[ -\frac{1}{2} \left( \frac{x - U_m}{\sigma} \right)^2 \right] \, dx \quad (3.25)
\]

where

- \( U_m \) is the mean value
- \( \sigma \) is the standard deviation

By defining a normalized variable as \( x_n = (x - U_m)/\sigma \), the expressions of the probability density and cumulative probability are simplified to

\[
f(u_n) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} x_n^2 \right) \quad (3.26)
\]
$$P(u_n) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u_n} \exp\left(-\frac{1}{2} x_n^2\right) dx_n$$  

(3.27)

Figure 3.5 illustrates the variations of the probability density and the cumulative probability as a function of the normalized variable $u_n$, expressed in units of $\sigma$, the standard deviation. One can verify the following properties of the normal distribution:

- Most probable voltage: $U_p = U_m$.
- Median voltage: $U_{50\%} = U_m$.
- The probability density is symmetrical with respect to $U_m$.

Parameters $\mu, \sigma^2$

Notation $N(m; \sigma^2) = F(x,m; \sigma^2)$

Expectation $EX = \mu$  

(3.28)

Variance $D^2X = \sigma^2$  

(3.29)

Point estimate

For $\mu$ : $\mu = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i = X_{50}$  

(3.30)

For $\sigma^2$ : $\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 = (X_{50} - X_{16})^2$  

(3.31)

where $X_{50}$ and $X_{16}$ are quantiles of the experimental distribution containing $n$ measurements.
Confidence estimate for $\mu$ can be defined from the $t$ distribution, and for $\sigma^2$ from the $\chi^2$ distribution, respectively

- For $\mu$: one-sided with upper limits, $I [−\infty, g_0]$, and lower limit, $I [g_1, +\infty]$

  \[
  \text{with } g_0 = X + t_{n-1,\epsilon} \frac{\sigma}{\sqrt{n}} \text{ and } g_1 = X - t_{n-1,\epsilon} \frac{\sigma}{\sqrt{n}}
  \]

  Two-sided limits: $I [g_1, g_0]$

  \[
  \text{with } g_0 = X + t_{n-1,1-\epsilon/2} \frac{\sigma}{\sqrt{n}} \text{ and } g_1 = X - t_{n-1,1-\epsilon/2} \frac{\sigma}{\sqrt{n}}
  \]

- For $\sigma^2$: one-sided with upper limits, $I [−\infty, g_0]$, and lower limit, $I [g_1, +\infty]$

  \[
  \text{with } g_0 = \frac{(n-1)\sigma^2}{\chi^2_{n-1,\epsilon}} \text{ and } g_1 = \frac{(n-1)\sigma^2}{\chi^2_{n-1,1-\epsilon}}
  \]

  Two-sided limits: $I [g_1, g_0]$

  \[
  \text{with } g_0 = \frac{(n-1)\sigma^2}{\chi^2_{n-1,1-\epsilon/2}} \text{ and } g_1 = \frac{(n-1)\sigma^2}{\chi^2_{n-1,1-\epsilon/2}}
  \]

### 3.2.3.2.3 Extreme Value Distribution

Introduced by Gumbel E.J. in his study on flooding, this distribution function is suitable for describing phenomena having large return periods. The probability density, $f(u)$, and the cumulative probability, $P(u)$, of the extreme value distribution are defined according to (Gumbel, 1954)

\[
f(u) = \frac{1}{\alpha} \exp\left[-\exp\left(\frac{u-U_0}{\alpha}\right)\right] \exp\left(\frac{u-U_0}{\alpha}\right) \quad (3.32)
\]

\[
P(u) = 1 - \exp\left[-\exp\left(\frac{u-U_0}{\alpha}\right)\right] \quad (3.33)
\]

where
- $U_0$ is the position parameter
- $\alpha$ is the form parameter

By defining a normalized variable as $u_n = (u - U_0)/\alpha$, the expressions of the probability density and cumulative probability are simplified to

\[
f(u_n) = \exp\left[-\exp(u_n)\right] \exp(u_n) \quad (3.34)
\]

\[
P(u_n) = 1 - \exp[-\exp(u_n)] \quad (3.35)
\]

Figure 3.6 illustrates the variations of the probability density and the cumulative probability as a function of the normalized variable, $(u - U_0)/\alpha$. One can verify the following properties of the extreme value distribution:

- At $u = U_0$, $P(U_0) = 0.632121$.
- Most probable voltage: $U_p = U_0$. 

---

---
• Median voltage: $U_{50\%} = U_0 + \alpha \ln(\ln 2)$.

• The density function is not symmetrical but slightly skewed at low probabilities ($U < U_0$).

Parameters $U_0$ and $\alpha$

Expectation $\text{EX} = U_0 - \alpha \lambda$ \hspace{1cm} (3.36)

with $\lambda = 0.5772$, Euler’s constant.

Variance $D^2X = \frac{1}{6} \pi^2 \alpha^2$ \hspace{1cm} (3.37)

Point estimate $U_0 = X_{63}$ and $\alpha = \frac{1}{3} (X_{63} - X_{05})$ \hspace{1cm} (3.38)

3.2.3.2.4 Weibull Distribution

Introduced by Weibull W., this distribution function is particularly suitable to describe phenomena with a threshold value, $U_0$, below which the probability of occurrence is identically zero. The probability density, $f(u)$, and the cumulative probability, $P(u)$, of the Weibull distribution are defined according to (Weibull, 1951)

$$f(u) = \frac{\beta}{\alpha} \left( \frac{u-U_0}{\alpha} \right)^{\beta-1} \exp \left[ -\left( \frac{u-U_0}{\alpha} \right)^{\beta} \right]$$ \hspace{1cm} (3.39)

$$P(u) = 1 - \exp \left[ -\left( \frac{u-U_0}{\alpha} \right)^{\beta} \right]$$ \hspace{1cm} (3.40)

where

$U_0$ is the position parameter
$\alpha$ is the scale parameter
$\beta$ is the form parameter

FIGURE 3.6 Extreme value distribution.
By defining a normalized variable as $u_n = (u - U_0)/\alpha$, the expressions of the probability density and cumulative probability are simplified to

$$f(u_n) = \beta (u_n)^{\beta-1} \exp\left[-(u_n)^{\beta}\right]$$  \hspace{1cm} (3.41)

$$P(u_n) = 1 - \exp\left[-(u_n)^{\beta}\right]$$  \hspace{1cm} (3.42)

Figure 3.7 illustrates the variations of the probability density and the cumulative probability as a function of the normalized variable, $u_n = (u - U_0)/\alpha$. One can verify the following properties of the Weibull distribution:

- There exists a threshold value $U_0$ such that the probability of occurrence $P(u)$ is identically zero for values of $u \leq U_0$.
- Most probable voltage: $U_p = U_0 + \alpha$.
- Median voltage: $U_{50\%} = U_0 + \alpha \exp\left[\frac{\ln(\ln 2)}{\beta}\right]$.
- The scale parameter $\alpha$ is a measure of the dispersion of the distribution.
- The probability density function is not symmetrical.

Special case: Exponential distribution is characterized by $U_0 = 0$, $\beta = 1$, and $\alpha = 1/\lambda$.
Parameters $U_0$, $\alpha$, and $\beta$

Expectation $\mathbb{E}X = U_0 + \alpha \Gamma \left( \frac{1}{\beta} + 1 \right)$ \hfill (3.43)

where $\Gamma \left( \frac{1}{\beta} + 1 \right)$ is the Gamma function.

For the exponential distribution,

$$
\mathbb{E}X = \frac{1}{\lambda} \hfill (3.44)
$$

Variance $D^2X = \alpha^2 \left[ \Gamma \left( \frac{2}{\beta} + 1 \right) - \Gamma^2 \left( \frac{1}{\beta} + 1 \right) \right]$ \hfill (3.45)

For the exponential distribution,

$$
D^2X = \frac{1}{\lambda^2} \hfill (3.46)
$$

### 3.3 Estimation of the Population Distribution

One of the main applications of statistical functions is the definition of a statistical model to adequately represent the complete population from a given set of experimental data, termed as sample of the population. In more specific terms, since the experimental sample is known, its statistic parameters are readily defined, and the question becomes how reliable it is to take the sample’s statistic parameters as representatives of complete population. In addition, since the type of the population is not known, it is usually assumed and validated with the sample population. These questions are object of a number of statistical tests; the most significant will be discussed in the following:

- Randomness test
- Using the sample’s statistic parameters
- Distribution test
- $\mu$-Selection test

#### 3.3.1 Randomness Test

Statistical analysis relies on the assumption that the experimental sample is random, that is, the measured data are independent and identically distributed according to the distribution of the population. This condition can be verified by observing the variation in chronological order of the data. A random sample will show oscillations of the data around a median value, with no tendency to steadily increase or decrease. This can be visualized by graphically representing the data in chronological order of their occurrence.

#### 3.3.2 Using the Sample’s Statistic Parameters

Once a test sample has been obtained, a natural approach would be to consider the sample’s statistic parameters as representatives of the population statistic distribution. How reliable is this approach? Consider the case where different samples are drawn from the same population and their statistical parameters evaluated. One obtains then new populations for each of the statistic parameters.
3.3.2.1 Sample Means

The population of the sample means \( \bar{X} \) will have an expected value:

\[
E(\bar{X}) = E\left[ \frac{1}{n} \sum X_i \right] = \frac{1}{n} \left[ \sum E(X_i) \right] = \mu
\]

The mean value of the population of sample means \( \bar{X} \) is also the mean value of the population \( x \). Noting that the independence of the sample data permits taking the sum of the variances for the variance of the sum, its variance is

\[
\text{var}(\bar{X}) = \text{var}\left( \frac{1}{n} \sum X_i \right) = \frac{1}{n} \left[ \sum \text{var}(X_i) \right] = \frac{\sigma^2}{n}
\]

The variance of the sample means decreases with \( n \) and approaches zero for large values of \( n \). In other words, the mean of a large sample of experimental data is representative of the population mean.

3.3.2.2 Sample Variance

The sample variance \( \sigma^2 \) is defined according to

\[
(n-1)\sigma^2 = \sum (X - \bar{X})^2
\]

which can be rewritten as

\[
(n-1)\sigma^2 = \sum (X - \mu)^2 - n(\bar{X} - \mu)^2
\]

Taking the expected value on both sides yields

\[
(n-1)E(\sigma^2) = \sum E[(X - \mu)^2] - nE[(\bar{X} - \mu)^2]
\]

\[
= \sum \sigma^2 - n \frac{\sigma^2}{n} = (n-1)\sigma^2
\]

and hence, the expected value of the sample variance is the same as the variance of the population \( x \):

\[
E(\sigma^2) = \sigma^2 \quad (3.48)
\]

The variance of \( \sigma^2 \) can be shown to be (Lindgren and McElrath, 1959)

\[
\text{var}(\sigma^2) = \frac{n}{n-1} \left[ E\left[ (X - \mu)^2 - \sigma^2 \right] \right]
\]

which will approach zero for large values of \( n \). The variance of a large sample is representative of the population variance.

The statistic parameters of the experimental sample will be representatives of the population of \( x \) when its size \( n \) is sufficiently large. Figure 3.8 shows the influence of the sample size on its distribution. One can see that as the sample size increases, the distribution becomes narrower and little improvement is noticed between samples of sizes between \( n = 50 \) and \( n = 100 \). A practical approach to define the sample size is to have its cumulative probability distribution cover an adequate range of probabilities. Typically,

- \( n = 10 \), the range of relative frequencies is \( P_{\text{min}} = 9.09\% \) to \( P_{\text{max}} = 90.91\% \).
- \( n = 100 \), the range of relative frequencies is \( P_{\text{min}} = 0.990\% \) to \( P_{\text{max}} = 99.01\% \).
3.3.3 Confidence Interval

In order to obtain a measure of the precision of an estimate \( \alpha^* \) of an unknown parameter \( \alpha \), we try to find two positive parameters \( \delta \) and \( \varepsilon \) such that the probability that the parameter \( a \) is included between the limits \( \alpha^* \pm \delta \) is equal to \( 1 - \varepsilon \):

\[
P(\alpha^* - \delta < \alpha < \alpha^* + \delta) = 1 - \varepsilon
\]

The range \( (\alpha^* - \delta, \alpha^* + \delta) \) is known as the confidence interval. The parameter \( 1 - \varepsilon \) is the confidence coefficient of the interval. And \( 100(1 - \varepsilon) \) may be termed the degree of confidence.

3.3.3.1 One-Sided Test

The statistical null hypothesis is

\[
H_0: \mu \leq \mu_0 \quad \text{against the alternative of } H_1: \mu > \mu_0
\]

The probability of accepting \( H_0 \), that is, the probability that the sample mean \( \bar{X} \) will not exceed \( \mu_0 \) in the case of large samples or for normal population, is

\[
P(\bar{X} < \mu_0) = 1 - F\left(\frac{\mu - \mu_0}{\sigma/\sqrt{n}}\right) = 1 - \frac{1}{\sigma\sqrt{2\pi n}} \int_{-\infty}^{\frac{\mu}{\sigma/\sqrt{n}}} \exp\left[-\frac{1}{2}\left(\frac{x - \mu_0}{\sigma/\sqrt{n}}\right)^2\right] dx
\]

(3.49)

Noting that Equation 3.49 is a normal distribution curve reversed, one can verify the effect of \( n \) and \( \mu_0 \) on the characteristic function defined by (3.49):

- Changing \( \mu_0 \) will shift the characteristic curve horizontally.
- The sample size determines the sampling variability in \( \bar{X} \): an increase in \( n \) will steepen the characteristic curve around \( \mu_0 \), approaching a unit impulse for large values of \( n \).
By specifying an accepted level $\alpha$, taken as the probability of rejection when $H_0$ is true,

$$\alpha = P(\text{reject } H_0) = P(\bar{X} \geq \mu_0) = F\left(\frac{\mu - \mu_0}{\sigma / \sqrt{n}}\right) = \frac{1}{\sqrt{2\pi}/n} \int_{\mu}^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \left(\frac{x - \mu_0}{\sigma / \sqrt{n}}\right)^2\right] dx$$

a critical region will be defined for rejection of $\mu_0$, the population mean, as illustrated in Figure 3.9a.

### 3.3.3.2 Two-Sided Test

If the statistic null hypothesis is

$$H_0: \mu = \mu_0 \quad \text{against the alternative of} \quad H_1: \mu \neq \mu_0$$

In this case, the critical region for rejecting $H_0$ is $|\bar{X} - \mu_0| > K$.

For a large sample or for a normal population, the probability of accepting $H_0$ is

$$P(|\bar{X} - \mu_0| < K) = F\left(\frac{\mu_0 + K - \mu}{\sigma / \sqrt{n}}\right) - F\left(\frac{\mu_0 - K - \mu}{\sigma / \sqrt{n}}\right)$$

(3.50)

considered as a function of $\mu$ for a fixed critical region and fixed sample size. By defining the significance level $\alpha$, taken as the probability of rejection when $H_0$ is true, a critical region will be defined for rejection of $H_0$, which is composed of two parts as illustrated in Figure 3.9b.

$$\alpha = P(\text{reject } H_0) = P(|\bar{X} - \mu_0| > K) = 2 \left[1 - F\left(\frac{K}{\sigma / \sqrt{n}}\right)\right] = 2F\left(-\frac{K}{\sigma / \sqrt{n}}\right)$$

(3.51)

From Equations 3.50 and 3.51, one gets the following relation:

$$P\left(|\bar{X} - \mu| < K\right) = 1 - \alpha$$

which can be rewritten as

$$P(\bar{X} - K < \mu < \bar{X} + K) = 1 - \alpha$$

(3.52)

which states that the inequality $(\bar{X} - K < \mu < \bar{X} + K)$ has a probability of $(1 - \alpha)$. However, this, by definition, is the $(1 - \alpha)$ confidence interval for the population mean $\mu$ as shown in Figure 3.9b.
3.3.4 Distribution Test

A population of experimental data does not, in general, have a specific type of distribution. Therefore, common practice assumes a type of distribution function, and its adequacy to represent the experimental set of data is statistically tested.

3.3.4.1 Chi-Square Test

It is the oldest test for goodness of fit still in use today. It allows evaluation whether an assumed population distribution or distribution type is consistent with the data of a random sample. The null hypothesis in this case is as follows:

- For a continuous variable: \( H_0 \) —the population density function is \( f(x) \).
- For a discrete variable: \( H_0 \) —the population probability function is \( p(x) \).

The alternative hypothesis \( H_1 \) is the population density or the probability function is not the one considered in \( H_0 \).

Consider a random sample of size \( n \); the frequency of the value \( x_i \) in the sample is \( p_{xi} \). If the population is discrete and defined by a density probability function \( f(x) = P(X = x_i) \), then the expected value of \( p_{xi} \) is \( n f(x_i) \). The deviations from the population density function can be expressed by a statistical variable \( \chi^2 \) defined as

\[
\chi^2 = \sum_{i=1}^{k} \frac{(p_{xi} - nf(x_i))^2}{nf(x_i)}
\]

It can be shown that for large samples, the variable \( \chi^2 \) has approximately the chi-square distribution with \((k - 1)\) degrees of freedom, \( k \) being the number of distinct possible values. The rejection region is taken to correspond to \( \chi^2 > k \). By selecting a significance level \( \alpha \), taken as the probability for rejecting \( H_0 \), a critical region for rejection of \( H_0 \) can be determined from the tabulated table of the chi-square function. A decision on the adequacy of the representation of the population distribution function by the random sample can be taken based on the value of \( \chi^2 \) with respect to the rejection limit.

3.3.4.2 Graphical Test

It is another way of evaluating the adequacy of a distribution function to represent a random sample. It has the advantage of being relatively simple and offers a visual evaluation of the adequacy of the model. The graphical test is based on the linearization of a typical distribution by the use of appropriate coordinate scales. An experimental distribution of the same type will be represented by a straight line, whose characteristics determine the parameters of the statistical model, the most representative of the experimental distribution. It should be noted that the graphical solution will require special graphic paper, one for each distribution function. The graphical linearization of a statistical distribution function, \( y = P(u) \), transforms the coordinates \( y \) and \( u \) into new variables \( z \) and \( v \), in a way that makes the relation between \( z \) and \( v \) linear (Figures 3.10 through 3.12) according to

\[
z = av + b
\]

where \( a \) and \( b \) are characteristic constants defining the straight line.

Let \( z = g(y) \) and \( v = h(u) \) be the transform functions. A graphical linearization of the function \( y = P(u) \) corresponds then to defining these transform functions. By replacing these transform functions in Equation 3.54, one gets

\[
z = g(y) = g\left[P(u)\right]
\]
Application of Statistical Analysis in High Voltage Engineering

For the three statistical distributions discussed, one can observe that

- The normalized variable \( u_n \) is given by \( u_n = P^{-1}(y) \)
- A linear relation already exists between \( u_n \) and \( u \) according to \( u_n = (u - U_0)/\alpha \)

One can show that selecting the transform functions \( g(y) \) and \( h(u) \) to be

\[
z = g(y) = P^{-1}(y) = u_n
\]

and

\[
v = h(u) = u
\]

FIGURE 3.10 Graphical representation of the normal distribution.

Equations 3.55 and 3.56 form the basis for defining graphical paper sheets, which are extensively used in statistical analysis; the most widely used graphical papers will be discussed in the following sections.
3.3.4.2.1 Normal Graphical Paper

When the distribution function is of the normal type, the transform functions are

$$
z = f^{-1} \left[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} \exp \left( -\frac{1}{2} x^2 \right) \, dx \right] = u_n \quad (3.57a)$$

and

$$v = u \quad (3.57b)$$

where $f^{-1}$ is the inverse function of the distribution $f$.

Referring to Figure 3.10, the linear relation describing the normal distribution that best fits the experimental distribution is

$$z = u_n = \frac{u - U_m}{\sigma} = \frac{1}{\sigma} \left( u - U_m \right) = \alpha v + b$$

The following observations can also be made:

- At $P = 0.5$, $v = 0$, $u = U_m$
- At $P = 0.8413$, $v = 1$, $u = U_m + \sigma$
- At $P = 0.1587$, $v = -1$, $u = U_m - \sigma$

**FIGURE 3.11** Graphical representation of the extreme value distribution.
This allows a quick evaluation of the characteristic parameters of a normal distribution for a given set of experimental data.

3.3.4.2.2 Extreme Value Graphical Paper
When the distribution function is of the extreme value type, the transform functions are
\[ z = g^{-1}(y) = \ln[-\ln(1 - y)] = u_n \]  
(3.58a)

and
\[ v = u \]  
(3.58b)

Referring to Figure 3.11, the linear relation describing the extreme value distribution that best fits the experimental distribution is
\[ z = u_n = \frac{u - U_0}{\alpha} = \frac{1}{\alpha}u - \frac{U_0}{\alpha} = av + b \]

And the following observations can be made:

- At \( P = 0.6321 \), \( v = 0 \), \( u = U_0 \)
- At \( P = 0.9340 \), \( v = 1 \), \( u - U_0 = \alpha \)
- At \( P = 0.3078 \), \( v = -1 \), \( U_0 - u = \alpha \)
This allows a quick evaluation of the characteristic parameters of an extreme value distribution representing a given set of experimental data.

### 3.3.4.2.3 Weibull Graphical Paper

When the distribution function is of the Weibull type, the transform functions are

\[
z = g^{-1}(y) = \exp\left(\frac{1}{\beta} \ln[-\ln(1-y)]\right) = u_n
\]  

(3.59a)

and

\[
v = u
\]  

(3.59b)

Referring to Equation 3.40, it can be seen that the Weibull distribution is characterized by three parameters: \(U_0\), \(\alpha\), and \(\beta\). As a result, there is no universal representation for this type of distribution function on graphical paper (which is necessarily bidimensional). Any graphical presentation of the Weibull distribution implies that at least one of the three parameters is known. Two particular cases will be discussed in the following, as related to the cases where the form parameter, \(\beta\), or the position parameter, \(U_0\), is known.

**Form parameter \(\beta\) is known:** When the form parameter \(\beta\) is known, the transposing functions are defined according to Equation 3.59, the Weibull distribution is represented by a straight line when the probability scale is defined according to Equation 3.59a, and the scale for the independent variable is linear according to Equation 3.59b. The linear equation of the straight line is (Figure 3.12)

\[
z = u_n = \frac{u - U_0}{\alpha} = \frac{1}{\alpha} u - \frac{U_0}{\alpha} = av + b
\]

The following observations can also be made:

- At \(p = 0\), \(v = 0\), \(u = U_0\)
- At \(p = 0.6321\), \(v = 1\), \(u = U_0 + \alpha\)

which allows a quick evaluation of the characteristic parameters of a Weibull distribution representing a given set of experimental data, namely, the position and scale parameters, \(U_0\) and \(\alpha\), respectively. The main inconvenience of this representation is that the probability scale varies with the value of \(\beta\), which means that a different graphical paper will be required for each value of \(\beta\).

**Position parameter \(U_0\) is known:** If the position parameter \(U_0\) is known, a more convenient representation of the Weibull distribution is obtained by taking the logarithm of \(z\). One then gets (Figure 3.13)

\[
\ln z = \frac{1}{\beta} \ln[-\ln(1-y)] = \ln(u_n)
\]  

(3.60a)

and

\[
v = \ln\left(\frac{u - U_0}{\alpha}\right) = \ln(u - U_0) - \ln \alpha
\]  

(3.60b)

The Weibull distribution is again represented by a straight line when the probability scale is according to Equation 3.60a, and the scale for the independent variable is logarithmic according to Equation 3.60b. This linearization of the Weibull distribution assumes that the position parameter \(U_0\) is known, and the linear equation of the straight line is

\[
\ln[-\ln(1-y)] = \beta \ln(u_n) = \beta \ln\left(\frac{u - U_0}{\alpha}\right) = \beta \ln(u - U_0) - \beta \ln \alpha = av + b
\]
which allows a quick evaluation of the characteristic parameters of a Weibull distribution representing a given set of experimental data, namely, the scale and form parameters $\alpha$ and $\beta$, respectively. In Figure 3.13b, the independent variable is normalized with respect to $U_0$ and $\alpha$; the point corresponding to the probable value ($v = 1, P = 0.6321$) is common to all distributions represented, while the slope of the straight lines is proportional to the form parameter $\beta$. It should be pointed out here that the independent variable is not $u$, but $(u - U_0)$. Only in the particular case corresponding to $U_0 = 0$, often used and referred to in practice as two-parameter Weibull distribution, the independent variable, $u$, is represented directly on a logarithmic scale.

**3.3.4.3 Linear Regression Analysis**

The linearization of the distribution function by means of appropriate coordinate systems allows the use of a powerful tool of statistical analysis, linear regression. This approach is particularly interesting with the widespread use of programmable calculators and personal computers. Indeed, since the graphical representation of the transposed distribution is a straight line,

$$z = av + b$$

linear regression can be put to good use to determine the equation of the straight line that is most representative of the experimental distribution. Using linear regression, we should obtain the parameters $a$ and $b$ according to (Miller and Freund, 1965)

$$a = \frac{\sum v_i z_i - \frac{\sum v_i}{n} \sum z_i}{\sum v_i^2 - \frac{(\sum v_i)^2}{n}} \quad \text{and} \quad b = \frac{\sum z_i}{n} - a \frac{\sum v_i}{n}$$

(3.61)
The quality of the correspondence of the straight line in (i.e., how good a fit we get) representing the experimental distribution is described by the correlation coefficient \( r^2 \) that is given by

\[
 r^2 = \frac{\left[ \sum v_i z_i - \frac{\sum v_i \sum z_i}{n} \right]^2}{\frac{\sum v_i^2}{n} \left( \frac{\sum z_i^2}{n} \right)} \tag{3.62}
\]

Regression analysis has the advantage of allowing a rapid and objective comparison of different statistical functions to represent a given set of experimental data.

### 3.3.4.3.1 Normal Distribution

When the distribution function is of the normal type, the transform functions are according to Equation 3.57. The characteristic parameters are derived from

\[
 U_m = -\frac{b}{a} \quad \text{and} \quad \sigma = \frac{1}{a} \tag{3.63}
\]

### 3.3.4.3.2 Extreme Value Distribution

When the distribution function is of the extreme value type, the transposing functions are according to Equation 3.58. The characteristic parameters are derived from

\[
 U_0 = -\frac{b}{a} \quad \text{and} \quad \alpha = \frac{1}{a} \tag{3.64}
\]

### 3.3.4.3.3 Weibull Distribution

When the distribution function is of the Weibull type, the transposing functions are according to Equation 3.58. The form parameter \( \beta \) is known and the other characteristic parameters are derived from

\[
 U_0 = -\frac{b}{a} \quad \text{and} \quad \alpha = \frac{1}{a} \tag{3.65}
\]

On the other hand, when the transform functions are according to Equations 3.60a and b, the position parameter \( U_0 \) is known, and the other characteristic parameters are derived from

\[
 \beta = a \quad \text{and} \quad \alpha = \exp\left( -\frac{b}{a} \right) \tag{3.66}
\]

### 3.3.5 Adequacy Representation of a Statistical Model

In general, since statistical distribution functions are not dictated by the phenomenon evaluated, they are useful as long as they adequately describe the phenomenon investigated. Furthermore, several statistical models may adequately represent the same set of data. As a result, it is of great interest to define the most suitable statistical distribution for a given set of data. This is particularly true when there is a sufficiently large amount of data available; the adequacy of statistical functions to represent the results may determine the choice of a particular distribution function. Regression analysis then provides a simple means to compare the quality of the representation of the experimental data by different statistical distribution functions.

However, most of the times, the limited amount of data available does not allow a definite differentiation between the models. The choice of a particular statistical function is often arbitrary, dictated by how familiar the user is with the use of the statistical functions. This is why the familiar normal distribution is the most often used in statistical analysis of experimental data. Sometimes, a good understanding of
the phenomenon evaluated can guide the choice of a statistical function. For example, one would use the
Weibull distribution to analyze a physical phenomenon that is known to have a threshold value \( U_0 \), below
which the probability of occurrence is nil.

For the set of data in Figure 3.1, different distribution functions are evaluated with respect to their
accuracy in describing the experimental distribution. Their characteristic parameters are shown in
Table 3.1 and in Figures 3.14 through 3.17. One can see that although the three distribution functions
adequately describe the experimental distribution, the results allow finer discrimination among the
distribution functions: The Weibull distribution is the most representative, whereas the extreme value
distribution is the least representative. Because it is known that the breakdown voltage of an oil gap
exhibits a threshold field value, the Weibull distribution represents a better choice.

In this context, it is worthwhile to establish the correspondence between the different most familiar
distribution functions. This will help the reader select the most appropriate statistical function for a par-
ticular situation. Although a complete correspondence among the different distribution functions is not
possible over the whole range of the independent variable, practical experience shows that it is possible
to obtain adequate representations for a set of experimental data by different distribution functions. It is
only required that the variations of the independent variable stay within a reasonable range, a condition
often met in practice.

### Table 3.1

<table>
<thead>
<tr>
<th>Distribution</th>
<th>( U_m ) (( U_0 ))</th>
<th>Sigma (Alpha)</th>
<th>Beta</th>
<th>Coefficient ( r^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental distribution</td>
<td>4.449</td>
<td>0.779</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Normal</td>
<td>4.498 kV/mm</td>
<td>0.941 kV/mm</td>
<td>—</td>
<td>0.977</td>
</tr>
<tr>
<td>Extreme values</td>
<td>4.967</td>
<td>0.828</td>
<td>—</td>
<td>0.933</td>
</tr>
<tr>
<td>Weibull—known ( U_0 )</td>
<td>2.5</td>
<td>2.265</td>
<td>2.398</td>
<td>0.985</td>
</tr>
<tr>
<td>Weibull—known ( \beta )</td>
<td>3.133</td>
<td>3.47</td>
<td>3.6</td>
<td>0.99</td>
</tr>
</tbody>
</table>

**FIGURE 3.14** Representation of the experimental distribution of the ac breakdown field, \( E_c \), of Figure 3.2, with a normal function.
The problem is illustrated in the following text where we evaluate how adequately a normal distribution can be represented by an extreme value or Weibull distribution. More specifically, the exercise aims at defining the range of variations of the independent variable over which it is possible to represent the normal distribution by another distribution function, of the extreme value and Weibull types, and to define how adequate those statistical models are.

**FIGURE 3.15** Representation of the experimental distribution of the ac breakdown field, $E_c$, of Figure 3.2, with extreme value function.

**FIGURE 3.16** Representation of the experimental distribution of the ac breakdown field, $E_c$, of Figure 3.2, with a Weibull function with known $U_0$. 
### 3.3.5.1 Extreme Value Distribution

Figure 3.18 compares the different approximations of the *normal* population by a distribution of the *extreme value* type. These distributions are selected in such a way as to intercept the *normal* distribution symmetrically with respect to the 50% probability level. Let $U_1$, $U_2$, $P_1$, and $P_2$ be, respectively, the voltages and probabilities at the interception points. The following relations may be derived:

$$ P_1 = 1 - \exp\left[-\exp\left(\frac{U_1 - U_0}{\alpha}\right)\right] \quad \text{and} \quad P_2 = 1 - \exp\left[-\exp\left(\frac{U_2 - U_0}{\alpha}\right)\right] $$

Therefore,

$$ \frac{U_1 - U_0}{\alpha} = \ln[-\ln(1 - P_1)] \quad \text{and} \quad \frac{U_2 - U_0}{\alpha} = \ln[-\ln(1 - P_2)] $$

It follows that

$$ \alpha = \frac{\ln[-\ln(1 - P_1)]}{U_1 - U_0} = \frac{\ln[-\ln(1 - P_2)]}{U_2 - U_0} \quad \text{(3.67)} $$

and

$$ U_0 = \frac{\ln[-\ln(1 - P_1)]U_1 - \ln[-\ln(1 - P_2)]U_2}{\ln[-\ln(1 - P_1)] - \ln[-\ln(1 - P_2)]} $$

Since the interception points are selected symmetrically with respect to the 50% probability,

$$ 1 - P_2 = P_1 $$
Equations 3.67 and 3.68 define the characteristic parameters of the extreme value distributions that best describe the normal distribution over the range of probabilities between \( P_1 \) and \( P_2 \). Table 3.2 compares the probability \( P(U_m) \) at the mean value \( U_m \) calculated for different extreme value distributions. It can be seen that as the interception points occur farther from the 50% probability level, the extreme value approximation becomes less adequate. However, the difference with the normal distribution, expressed by the ratio \( U_0/U_m \), remains reasonable, in the order of 15%, for the range of probabilities between 5% and 95%.

3.3.5.2 Weibull Distribution

The different Weibull distributions evaluated are selected in such a way as to have the same position parameter \( U_0 \) defined by

\[
U_0 = U_m - k\sigma
\]

<table>
<thead>
<tr>
<th>Probabilities at Interception Points</th>
<th>Probability at ( U_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.15</td>
<td>0.85</td>
</tr>
<tr>
<td>0.06</td>
<td>0.94</td>
</tr>
<tr>
<td>0.02</td>
<td>0.98</td>
</tr>
</tbody>
</table>

TABLE 3.2

Extreme Value Representation of a Normal Data Population

FIGURE 3.18 Adequacy of extreme value distributions to represent a normal population.
and adjusting the scale parameter, \( \alpha \), so that the 50% probability level coincides with the mean value of the normal distribution. By varying \( k \), the optimal position parameter \( U_0 \) was found to correspond to

\[
U_0 = U_m - 3.33\sigma
\]

The corresponding value of the scale parameter \( \alpha \) is

\[
\alpha = 3.69\sigma
\]

Figure 3.19 compares different Weibull distributions for different values of the form parameter \( \beta \). The best result is obtained with the value of the form parameter:

\[
\beta = 3.6
\]

which approximates the normal distribution to within 1% over the range of probabilities between 1% and 99%.

### 3.3.5.3 Discussion

From the preceding analysis, the following observations can be made:

- If accuracy of the order of 5% is sufficient, the three distribution functions are equivalent over the range of probabilities between 5% and 95%.
- To be able to differentiate the different statistical models, the experimental distribution must be defined over a wide range of probabilities, typically between \( 10^{-3} \) and 0.999, a requirement that may be difficult to meet in practice due to the generally high cost associated with the acquisition of experimental data.
• There exists a good correspondence between the normal and Weibull distributions, as the characteristic parameters of the latter may be adjusted to provide a good concordance with the normal distribution.

• Best results are obtained with $U_0 = U_m - 3.33 \sigma$, $\alpha = 0.69 \sigma$, and $\beta = 3.6$, which adequately represent a normal distribution over the range of probabilities from 1% to 99%.

3.3.5.3.1 International Electrotechnical Commission Procedure for Estimating a Weibull Distribution (IEC 6071-2, 1996)

This procedure determines the equivalent Weibull distribution to fit experimental data which has been presented by a Normal distribution ($U_{50}$, $\sigma$). It requires that the position parameter $U_0$ be estimated first, usually taken as the withstand voltage:

$$U_0 = U_{50\%} - n\sigma$$

where $n$ is usually taken as $n = 2, 3, 4$ and $\sigma$ the standard deviation.

The remaining two parameters, $\alpha$ and $\beta$, of the Weibull distribution can then be defined from the condition that the two distributions have the same values at $P = 0.5$ and $P - 0.16$, which results in the following expressions for $\beta$ and $\alpha$ of the equivalent Weibull distribution:

$$\beta = \frac{1.38}{\ln \left( \frac{n}{n-1} \right)} \quad \text{and} \quad \alpha = \frac{n\sigma}{(\ln 2)^{1/\beta}}$$

As a numerical example, for $n = 2$, $\beta = 2$ and $\alpha = 2.4 \sigma$ while for $n = 3$, $\beta = 3.4$ and $\alpha = 3.3 \sigma$.

3.3.6 Direct Extrapolation of Experimental Data

Once the statistical model has been defined for a set of data, the various characteristic parameters of the experimental distribution can readily be established. However, the main advantage of statistical functions resides in their ability to extrapolate the behavior of a physical system beyond the range covered by the experimental results. Consider, for example, experiments aimed at determining the breakdown probability of an insulation system. Whereas it is relatively easy to obtain experimental data for high-probability events (i.e., events that occur between 10% and 90% of the times), real-world insulation systems must be assessed for much less probable events, down to withstand probabilities of $10^{-3}$ or even lower. Experimental evaluation of such low breakdown probability distribution often turns out to be simply too expensive.

The usual practice relies on the definition of an appropriate statistical model that can then be used to extrapolate the results to those low-probability levels. The selection of the statistical model is of prime importance in this case since it can have a significant influence on the extrapolated results. By referring to the results of Figures 3.18 and 3.19, it is obvious that if adequate representations of the normal population can be obtained by the carefully selected extreme value and Weibull distributions over a range of relatively high probabilities, their extrapolations to the low-probability levels of $10^{-3}$ show considerable differences.

The best results are obtained when extrapolation is made in accordance with the physical phenomenon investigated to allow proper extrapolation to very-low-probability withstand levels. The Weibull distribution, for example, which has a threshold value, $U_0$, below which the probability of occurrence is zero, is particularly suitable to describe many insulation phenomena that usually have an onset field.

3.3.7 Verification of an Experimental Law

Experimental evaluation is often required to validate empirical or semiempirical laws describing the behavior of the physical system under investigation. This implies that experimental data must be obtained for different conditions, selected over a range of the independent variable. The results are presented by data points corresponding to the mean values, with the range of variation of the dependent variable indicated by a linear bar extending on both sides to indicate an experimental error. When sufficient experimental data are available, an experimental distribution and its confidence limits can be defined for each data point, defining the two boundaries within which most of the experimental data are expected to occur.
An example is given in Figure 3.20 that illustrates the experimental evaluation of thermal aging of paper in oil: The tests are made on different sample populations of paper, each corresponding to a different temperature. The useful life of the paper was defined as the aging time taken to reach 20% of its initial tensile strength. The results are used to determine the parameters defining the thermal aging process, according to the Arrhenius model:

\[ T_b = A \exp \left( \frac{k}{T} \right) \]  

(3.69)

where
- the thermal life \( T_b \) is the time to breakdown at constant temperature
- \( T \) is the temperature in degree K
- the constant \( A \) is characteristic of the paper aging process

Taking the logarithm of both sides, the preceding relation becomes

\[ \ln(T_b) = \frac{k}{T} + \ln A \]  

(3.70)

Equation 3.70a indicates that the thermal life function will be a straight line when plotted on a graphical paper with \( \ln(T_b) \) in ordinates and \( 1/T \) in abscissas. The constant \( A \) is then equal to the slope of the experimental function. In Figure 3.20, each sample population is described by its mean and the 90% confidence limits estimated from the experimental sample data. The straight line corresponding to the sample means validates the Arrhenius law, while the confidence region corresponding to the 90% confidence limits represents the adequacy of the experimental data to describe the theoretical law.

Since it is economically unjustified to conduct aging tests at stresses comparable to the operating value, they are generally made under conditions that are significantly more severe than those encountered in operating conditions of the equipment. The results are extrapolated to service conditions following an empirical model. Extrapolation of accelerated aging results assumes that the aging mechanism...
remains the same under accelerated aging test as in actual operating conditions, a condition not always met in practice. Furthermore, several aging mechanisms may be active at the same time, which further complicate the interpretation of aging test results.

3.4 Weak Link Statistics

Real insulation systems, because of their size and complexity, are not subjected to extensive dielectric tests. Instead, a typical section or individual components of the system can be evaluated in depth in laboratory and their dielectric performances extrapolated to the real system. Such an extrapolation process assumes that

- The real insulation system is formed by \( n \) identical components
- The performance of the individual components are well defined (i.e., there is enough experimental data to derive an adequate statistical model)
- The breakdown phenomenon is the same in laboratory as it is for the system under operating conditions

Weak link statistics are particularly useful to extrapolate the experimental results in this case. Let \( P_i(u) \) be the cumulative breakdown probability of a single component at the voltage \( u \). The withstand probability \( Q_i(u) \) of that component is given by

\[
Q_i(u) = 1 - P_i(u)
\]  

(3.71a)

The withstand of \( n \) identical components connected in parallel can be interpreted to mean that each individual component can withstand the test voltage \( u \). Consequently, the withstand probability of the complete system can be expressed as

\[
Q(u) = \prod_{i=1}^{n} Q_i(u) = [1 - P_i(u)]^n
\]  

(3.71b)

It follows that the breakdown probability of the complete system of \( n \) parallel components is

\[
P(u) = 1 - Q(u) = 1 - \prod_{i=1}^{n} Q_i(u) = 1 - [1 - P_i(u)]^n
\]  

(3.72)

with \( P_i(u) \) assumed independent.

The breakdown probability of the whole system is readily evaluated from the breakdown probability distribution of the elementary gap. Weak link statistics are particularly useful as they take into account the length effect of high-voltage cables, discussed in Chapter 11, as well as the area and volume effects observed in the breakdown of oil gaps, discussed in Chapter 12. Weak link statistics are also useful in the evaluation of the statistical significance of dielectric test methods, especially when low-probability withstands are required. This last topic will be discussed in Section 3.7.2.

3.5 Optimization of Experimental Data

Until now, the statistical models have assumed that the experimental data evaluated formed one unique and homogeneous population that could be described by a single distribution function. Such a situation does not always occur in practice. Typical real-world experimental data may be made out of several component populations. An indication of this situation is when the experimental distribution cannot be adequately described by a single distribution function. Additionally, the shape of the cumulative distribution of probability will frequently show one or several pronounced inflections.
3.5.1 Determination of the Component Distributions

When there are more than one component populations present, the resulting probability, \( P(x) \), can be expressed as a function of the component distributions according to

\[
P(x) = \sum k_i P_i(x) \quad \text{with} \quad \sum k_i = 1, \quad i = 1, n
\]

where

- \( k_i \) represents the contribution \( P_i(x) \) of the \( i \)th component population to the resulting probability
- \( n \) is the number of component distributions

The number of component distributions can be determined from the graphical representation of the experimental distribution as \( n = l + 1 \), \( l \) being the number of inflection points identified in the resulting distribution. The solution of the previously presented system of equations defines completely the \( n \) component distributions of the experimental data. The straightforward way to determine the component distributions consists in deriving a system of \( m \) equations for the \( m \) characteristic parameters, defining the \( n \) component distributions according to

\[
P(x) = \sum k_j P_j(a_j, x), \quad \text{with} \quad j = 1, m
\]

where \( a_j \) represents the characteristic parameters that define the component distributions \( p_j(x) \).

A simplified approach can also be used by subdividing the experimental data into \( n \) component populations according to the coordinates of the inflection points. Linear regression analysis can then be used to define the appropriate distribution for each of the component populations. In this approach, the component distributions are directly affected by the subdivision of the experimental data into component populations. By trial and error and using the coefficient of correlation as an objective criterion of selection, the component populations can be gradually adjusted to determine the best representation for the complete set of experimental data. This approach has the advantage of using regression analysis that is relatively simple and rapid to perform. However, knowing the component distributions embedded in a set of data does not by itself provide any explanation of the phenomenon investigated. It does give a better, more coherent view of the data available and, in turn, may help interpret them. The few example cases in the following can illustrate this point.

3.5.1.1 Audible Noise

An example of multiple populations is shown in Figure 3.21, which shows the audible noise (AN) performance of a conductor bundle considered for HVDC power transmission at ±750 kV dc. The AN measured at 15 m from the positive pole was recorded over a period of 12 months during the corona test at IREQ. It can be seen that the distribution of the AN presents two inflection points, which implies that the data are composed of three component populations.

Table 3.3 gives the characteristic parameters of the component distributions for the various seasonal distributions of the AN recorded. Two of these component distributions are related to the corona activities under fair and foul weathers, while the third component distribution was ultimately identified to match the local traffic noise. After removing the traffic noise component distribution, a better interpretation of the remaining experimental data was possible as illustrated by the calculated annual distribution in Figure 3.21.

3.5.1.2 Streamer Length

In the study of discharges in oil, a series of measurements of the streamer lengths was made during a dielectric withstand test (Figure 3.22). The cumulative distribution shows one inflection point, hinting at the presence of two distinct populations in the set of data. Table 3.4 gives the characteristic parameters of the component distributions. Indeed the data set can be broken in two distinct groups, one characterized by frequent and short streamers and a second group of longer streamers. It is apparent that a more
vigoroues process is present, which generates breakdown streamers that extend further into the oil gap eventually causing its breakdown.

### 3.5.2 Improvement of Experimental Data

Analysis of experimental data is often limited by the small amount of data available. One can get around this problem by defining new variables that allow grouping data from different tests into a

<table>
<thead>
<tr>
<th>Distribution #</th>
<th>Mean Value</th>
<th>Standard Deviation</th>
<th>Percentage %</th>
<th>Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>50.3</td>
<td>1.2</td>
<td>87</td>
<td>Fair weather</td>
</tr>
<tr>
<td>2</td>
<td>47.0</td>
<td>1.1</td>
<td>13</td>
<td>Foul weather</td>
</tr>
<tr>
<td>Autumn</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>47.0</td>
<td>1.9</td>
<td>85.5</td>
<td>Fair weather</td>
</tr>
<tr>
<td>2</td>
<td>44.2</td>
<td>2</td>
<td>10</td>
<td>Foul weather</td>
</tr>
<tr>
<td>3</td>
<td>57.1</td>
<td>1.5</td>
<td>4.5</td>
<td>Traffic</td>
</tr>
<tr>
<td>Winter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>48.5</td>
<td>2.2</td>
<td>74.6</td>
<td>Fair weather</td>
</tr>
<tr>
<td>2</td>
<td>41.8</td>
<td>1.5</td>
<td>20.6</td>
<td>Foul weather</td>
</tr>
<tr>
<td>3</td>
<td>59.4</td>
<td>0.7</td>
<td>4.8</td>
<td>Traffic</td>
</tr>
<tr>
<td>Annual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>48.2</td>
<td>1.8</td>
<td>83.2</td>
<td>Fair weather</td>
</tr>
<tr>
<td>2</td>
<td>44.3</td>
<td>1.65</td>
<td>13.4</td>
<td>Foul weather</td>
</tr>
</tbody>
</table>

larger population, more appropriate for statistical analysis. An example of regrouping data is in the evaluation of the withstand capability of gas-insulated substation (GIS) to internal arc conditions. Under conditions of stationary arc, the time to burn-through of the GIS envelope can be expressed by the functions of the envelope thickness and of the arc current, according to the following empirical relations (Trinh, 1992):

\[ T_p = k_i E_p^{\alpha} \quad \text{and} \quad T_p = \frac{k_i}{I_c^\beta} \]

where
- \( T_p \) is the time to burn-through in milliseconds
- \( E_p \) is the envelope thickness in millimeters
- \( I_c \) is the arc current in kA
- \( \alpha, \beta, k_i, \) and \( k_i \) are experimental constants

It has proven difficult to verify these behaviors when experiments were carried out under conditions approaching practical operating conditions of GIS. Despite the large amount of experimental data
gathered, the measured times to burn-through were widely dispersed due to the wide range of experimental conditions. By defining a new variable $C_p$ according to

$$C_p = \beta \frac{I_c}{E_p}$$  \hspace{1cm} (3.73)

the effect of all three variables, $T_p$, $E_p$, and $I_c$, can be combined and, as a result, allows the regrouping of all experimental data into a single and much larger set of data for the new variable $C_p$, as illustrated in Figure 3.23. The results show that the set of data for $C_p$ is formed of two distinct populations, implying two different burn-through processes. These are associated with stationary arcs with very short times to burn-through and mobile arcs with much longer time to burn-through and larger standard deviations associated with the traveling time before the arc becomes stationary and cause the burn-through.

### 3.5.2.1 Breakdown Probability of Elementary Volume

Another example of regrouping data is in the evaluation of the breakdown probability of an elementary volume in large volume oil insulation systems (Wilson, 1953). We saw earlier that Equation 3.72 described the breakdown probability of an oil gap comprised of $n$ unit volumes. We also saw that the breakdown probability of the unit volume, $P_i(u)$, is related to the breakdown probability of the complete oil gap, according to

$$P_i(u) = 1 - \sqrt[n]{1 - P(u)}$$  \hspace{1cm} (3.74)
As a result, the breakdown probability distribution of the unit oil volume can be evaluated over a wide range of low probabilities from the available experimental data obtained with oil gaps of different oil volumes, thus taking advantage of a much larger population than is possible with any single experiment.

### 3.6 Joint Probability

When the outcome of an event depends on two distinct phenomena, the probability of its occurrence can be readily evaluated as follows:

- Let \( f(u) \) be the marginal probability density distribution that the marginal variable will have a value of \( u \).
- Let \( P(u) \) be the cumulative conditional probability that the marginal variable \( u \) will cause the observed outcome.

The elementary joint probability \( dR \) is simply the product of the probability density of occurrence \( f(u) \) \( du \) of a marginal variable, \( u \), by the conditional cumulative probability \( P(u) \) that it will lead to the outcome of the corresponding event:

\[
dR = P(u) f(u) du
\]

The joint probability \( R \) is the sum of all the elementary joint probabilities evaluated over the whole range of the marginal variable:

\[
R = \int_0^\infty f(u) P(u) du \tag{3.75}
\]

#### 3.6.1 Risk of Failure

One application of the joint probability is in the evaluation of the risk of failure of a physical system, for example, a line tower, in the normal operating conditions of the line. In this case, \( f(u) \) is the marginal probability density that the overvoltages produced by the normal switching operations on the line will have a magnitude of \( u \), and \( P(u) \) is the conditional cumulative probability that the overvoltage \( u \) will cause breakdown of the tower clearances. Risk of failure is currently used to provide a quantitative evaluation of the quality of an insulation system, discussed in Chapter 13.

#### 3.6.2 \( V-t \) Breakdown Characteristic

When subjected to impulse voltage of very steep front, for example, under conditions of steep lightning pulses or during disconnect-switch operations in a SF\(_6\)-insulated substation, the breakdown of an insulation system tends to occur at voltage above the normal withstand level, following a very short time that the breakdown process needs to be initiated and to develop. Joint probability is used by Rizk and Eteiba to evaluate the \( V-t \) breakdown characteristic of an insulating gap (Rizk and Eteiba, 1982), discussed in Chapter 7. Here, \( f(u) \) is the marginal probability density that the impulse voltage will have a prospective peak of magnitude \( u \), and \( P(ult_d) \) is the conditional probability that the breakdown will occur after a delay time \( t_d \). By evaluating the joint probability of occurrence for various values of \( u \), a region in the \( V-t \) plane can be delimited between the two boundary curves of 5% and 95% probability where most of the breakdowns will occur.
3.7 Verification of Characteristics or Behavior by Specific Tests

Engineering work requires quantitative assessment of the system performance as well as objective characterization prior to their deployment in service. Specific tests are therefore needed to demonstrate the performance of the system. Statistical analysis provides the tools to help in planning the tests so as to minimize the cost and effort. The following aspects may be considered:

- Optimization of the number of test samples
- Statistical significance of test procedures
- Computer simulation for developing new test methods

3.7.1 Optimization of the Number of Test Samples

Referring to Section 3.2.2.1, the cumulative experimental distribution of a set of experimental data is defined over the range of probabilities comprised between

\[ P_{\text{min}} = \frac{1}{N+1} \quad \text{and} \quad P_{\text{max}} = \frac{N}{N+1} \]

3.7.1.1 Minimum Required Number of Tests

The earlier expressions of extreme probabilities provide a simple means to define the minimum number \( N \) of tests required to obtain either a given withstand level or an assured breakdown level associated, respectively, with \( P_{\text{min}} \) and \( P_{\text{max}} \):

\[ N = \frac{1}{P_{\text{min}} - 1} \quad \text{or} \quad N = \frac{1}{1 - P_{\text{max}}} \]

The minimum required number of tests can become very large rapidly as the required withstand level gets lower. This provides a motivation to develop special test methods for low-probability breakdowns and withstand voltages (Trinh and Vincent, 1980). For most practical cases, a probability distribution covering the range of probabilities between 10% and 90% is adequate. By using statistical analysis and extrapolation, the minimum number is \( n = 9 \) test samples.

It should be borne in mind that this minimum number of test samples will provide only one point on the performance curve corresponding to a specific set of test conditions. This number should be multiplied by the total number of test conditions considered, and the total number of test samples required may rapidly reach levels that become economically unjustified. Specific test procedures may be considered when the emphasis is on the determination of a specific performance, for example, a low-probability withstand level, which allows optimization of the number of test samples as well as the testing time and cost.

3.7.2 Statistical Significance of Test Procedures

Low-probability withstand tests aim at minimizing the number of tests, especially if they are destructive, while determining the withstand levels with very low breakdown probabilities. Most of these tests rely on designing the method of application of the test voltage, using two basic approaches corresponding to

- Test at constant voltage
- Test with increasing voltage

Weak link statistics can be used to evaluate the statistical significance of different test procedures for achieving the desired low-probability withstands (Trinh and Vincent, 1980).
3.7.2.1 Test with Impulse Voltage

When the dielectric performance of an insulating system is evaluated under impulse voltages, a breakdown probability function $P_i(u)$ can be associated to each voltage application of amplitude $u$.

3.7.2.1.1 Test at Constant Voltage

In this test procedure, the test voltage is kept at the voltage level $u$ and a total of $n$ voltage applications will be made to the test object. Since the breakdown probability of a single application of the voltage $u$ is $P_i(u)$, the corresponding withstand probability is

$$Q_i(u) = 1 - P_i(u)$$

The withstand probability of the insulation system after $n$ voltage applications is

$$Q_n(u) = \prod_{i=1}^{n} Q_i(u) = \left[ 1 - P_i(u) \right]^n$$

and the breakdown probability of the system after $n$ voltage application is

$$P_n(u) = 1 - Q_n(u) = 1 - \prod_{i=1}^{n} Q_i(u) = 1 - \left[ 1 - P_i(u) \right]^n$$ (3.76)

Figure 3.24 compares the cumulative breakdown probability of tests at constant voltage with different values of $n$, where $n$ is the number of voltage applications. The breakdown probability of a single voltage application $P_i(u)$ is assumed to be of normal type. As expected, the breakdown probability in a test at constant voltage increases with $n$. This means that tests at constant voltages can be performed at much

---

**FIGURE 3.24** Breakdown probability under constant impulse and ac voltage—normal distribution. (From Trinh, N.G. and Vincent, C., IEEE Trans. Power Appar. Syst., PAS-99, 711, 1980.)
lower voltages while assuring a measured low intrinsic probability withstand level. It also limits at the same time the number of test samples effectively damaged during the test series.

3.7.2.1.2 Test with Increasing Step Voltages (Rizk and Vincent, 1977)

In this test procedure, the test starts at a voltage $U_0$ with very low breakdown probability; the test voltage is then increased by equal steps $\Delta u$ with $n$ voltage applications at each level, until breakdown occurs at the voltage $u$. Let $m$ be the number of voltage steps to reach the test voltage $u$:

$$u = U_0 + m\Delta u$$

Let $P(u_i)$ be the breakdown probability of a test gap at the voltage level $u_i$; its withstand probability $Q(u_i)$ is

$$Q(u_i) = 1 - P(u_i)$$

since there are $n$ voltage applications at $u_i$, the withstand probability $Q(u_i)$ at the voltage level $u_i$ is

$$Q(u_i) = \left[ Q(u_i) \right]^n = \left[ 1 - P(u_i) \right]^n$$

The withstand of the test gap after $m$ voltage levels implies that the gap withstands all the voltage levels below $u$. Consequently, the withstand probability of the gap is

$$Q_n(u) = \prod_{i=1}^{m} Q(u_i) = \prod_{i=1}^{m} \left[ 1 - P(u_i) \right]^n$$

It follows that the breakdown probability of the test gap after $n$ voltage levels is

$$P_n(u) = 1 - Q_n(u) = 1 - \prod_{i=1}^{m} Q(u_i) = 1 - \prod_{i=1}^{m} \left[ 1 - P(u_i) \right]^n$$

(3.77)

and is lower than the breakdown probability of the gap at the voltage $u$. Figure 3.25 compares the cumulative breakdown probability of tests at increasing voltage with different values of $n$, the number of voltage applications. The breakdown probability of a single voltage application $P(u)$ is assumed to be of the Weibull type. As expected, the breakdown probability in a test with increasing voltage with $n$ voltage applications at each level increases even more than in test at constant voltages. This means that test with increasing voltages can be performed at still much lower voltages while assuring a measured low-probability withstand level. It also limits at the same time the number of test samples effectively damaged during the test series.

3.7.2.2 Practical Test Procedures

Most practical test procedures for low-probability withstands are a combination of the earlier two basic test procedures. Their statistical significance can therefore be readily evaluated from the knowledge of the breakdown probability under single voltage application.

3.7.2.2.1 IEC Rated Withstand Voltage Test (IEC 60060-1, 1989)

In this test, 15 impulses of rated withstand voltage with the specified shape and polarity are applied. The requirements of the test are satisfied if not more than two breakdowns occur. For a given intrinsic probability $p(u)$ of breakdown in one single voltage application $P(u)$ is assumed to be of the Weibull type. As expected, the breakdown probability in a test with increasing voltage with $n$ voltage applications at each level increases even more than in test at constant voltages. This means that test with increasing voltages can be performed at still much lower voltages while assuring a measured low-probability withstand level. It also limits at the same time the number of test samples effectively damaged during the test series.

$$P_{\text{test}} = P_0(k = 0, n = 15) + P_1(k = 1, n = 15) + P_2(k = 2, n = 15)$$

where $P_0$, $P_1$, and $P_2$ are, respectively, probabilities to have 0, 1, and 2 breakdowns in a series of 15 tests. For an intrinsic breakdown probability of $p = 0.05$, the binomial distribution probabilities $P_0 = 0.463$, $P_1 = 0.366$, and $P_2 = 0.135$ result in a probability $P_{\text{test}} = 0.964$ to pass the test.
Another interpretation of the IEC rated withstand test, from Equation 3.3, is to associate it to a withstand probability of $2/(15 + 1)$ or less in a test at constant voltage with 15 voltage applications. Referring to the results of Figure 3.26, the IEC rated withstand voltage test assures an intrinsic breakdown probability of $10^{-2}$ or less as currently accepted in practice.

### 3.7.2.2.2 ANSI Rated Withstand Voltage Test (ANSI, 1968)

A different test procedure is proposed by the ANSI for the rated withstand voltage, in two steps. Five impulses of rated withstand voltage with specified shape and polarity are first applied. If no breakdown occurs, the requirements for the test are satisfied. This corresponds to a withstand probability of $1/(5 + 1)$ in a test at constant voltage with 5 voltage applications. Referring to the results of Figure 3.26, this ANSI rated withstand voltage assures an intrinsic breakdown probability of a few percent.

If only 1 breakdown occurs, 10 additional impulses are then applied. The test requirements are again satisfied if no breakdown occurs in the additional 10 voltage applications. This corresponds to a withstand probability of $1/(10 + 1)$ in a test at constant voltage with 10 voltage applications, subjected to the condition that there is only 1 breakdown in the first series of 5 tests. Referring to the results of Figure 3.26, the corresponding intrinsic breakdown probability in this case is $10^{-2}$ or less.

### 3.7.2.2.3 Tezner Test Procedures (Tezner, 1958)

It is essentially a test with increasing voltage, with a single voltage application at each level, $n = 1$. Furthermore, the test can start at a voltage level $U_0$ with a finite probability of breakdown. Figure 3.26
illustrates the changes in the breakdown probability with Tezner test procedures with impulse voltages from the breakdown probability under single voltage application \( P_i(u) \), assumed to be of the extreme value type. Tezner test procedure assures an intrinsic breakdown probability in the range of 0.1–0.3.

### 3.7.2.3 Test with AC Voltage

In test with ac voltage, a similar evaluation of the statistical performance of test procedures is also possible by defining the 1-cycle breakdown probability, of the insulation system, and applying weak link statistics as in the case of testing with impulse voltages. Figures 3.24 through 3.26 illustrate also the changes in the breakdown probability with different test procedures from the one-cycle breakdown probability \( P_i(u) \), assumed to be of the normal, extreme value, or Weibull type.

#### 3.7.2.3.1 Rated AC Withstand Voltage Test

In this test, the ac voltage is raised to the rated withstand voltage, maintained at this level for the specified time and then rapidly decreased. The test is satisfied if no breakdown occurs. The currently accepted withstand period is 1 min. Referring to the results of Figure 3.24, the rated ac withstand voltage test assures an intrinsic breakdown probability in one cycle of the applied voltage amounts to \( 10^{-2} \) or less.
3.7.2.3.2 AC Test with Ramp Voltages

The corresponding Tezner test with AC voltage is obviously the case of testing with ramp voltages. If \( \nu \) is the rate of rise of the test voltage and \( f \) is the frequency, then the equivalent voltage step is

\[
\Delta u = \frac{\nu}{f}
\]

Since the ramp voltage always starts from zero, the measured critical flashover (CFO) voltage is a function of the rate of rise \( \nu \) of the test voltage. Typical intrinsic breakdown probability of the measured CFO voltage obtained with this test procedure is about \( 5 \times 10^{-3} \) to \( 2 \times 10^{-3} \) for a rate of rise \( \nu \) of 0.1–1.0 \( \sigma/s \), where \( \sigma \) is the estimated standard deviation of the breakdown voltage (Figure 3.27).

3.7.2.3.3 AC Test with Increasing Voltages

In AC test with increasing voltages, the test voltage is raised in steps \( \Delta u \) from a low initial voltage \( U_0 \). At each voltage level, the test voltage is kept constant for a time \( T \). If the time required for raising the test voltage is small compared to the withstand time \( T \), this test procedure is essentially a test with increasing voltage and \( n \) voltage applications at each level, \( n \) being equal to \( N = fT \).

To save testing time (and cost), a modified procedure is often used: the test voltage is raised rapidly to the first withstand level close to about 80% of the expected breakdown voltage and then in increasing steps of \( \Delta u \) until breakdown occurs. Referring to the results of Figure 3.26, a typical one-cycle intrinsic breakdown probability of the measured CFO voltage is in the range of \( 10^{-4} \) for AC test with increasing voltage and 1 min withstand periods.

**FIGURE 3.27** Correspondence between practical test procedures with AC voltages. (From Trinh, N.G. et al., *IEEE Trans. Power Appar. Syst.*, PAS-101, 3712, 1982.)
3.7.2.3.4 The 1 and 30 min Withstand Voltage of an Oil Gap

An application of this approach is in the evaluation of the 1 and 30 min. withstand voltages of an oil gap. The results of Figure 3.28 compare the statistical significance of the 1 and 30 min withstand test with the steadily rising voltage test for which the breakdown probability distribution of the unit volume was derived, by assuming a complete correspondence between different test procedures. It can be seen that the 50% withstand level in the 1 and 30 min withstands is equivalent to the $0.85 \times 10^{-2}$ and $1.8 \times 10^{-4}$ withstand levels in a test with steadily increasing voltage. As a result, the 1 and 30 min withstand voltages can be readily determined from the breakdown probability distribution curve of the unit volume derived from test results obtained in the test with steadily increasing voltage. Typical results are presented in Figure 3.28 that compares the calculated 1 and 30 min withstand levels with experimental measurements and illustrate the validity of this approach (Trinh et al., 1982).

3.7.3 Computer Simulation of Test Procedures

Weak link statistics provide a simple means of evaluating new test methods, through computer simulation (Rizk and Vincent, 1977). The required test sequence is programmed to the computer, starting with the selection of the test voltage $u$. The breakdown probability $P(u)$ is then calculated according to the test procedures and the statistical function specified. The decision regarding whether the voltage application results in a breakdown or withstand is made in two steps:

- Generate a pseudorandom number $k$ between 0 and 1.
- By comparing $k$ with the calculated cumulative breakdown probability $P(u)$, a breakdown occurs if $k \geq P(u)$, otherwise the gap withstands.
By repeating $N$ times the simulation and noting the simulated result of each computer simulation, the cumulative breakdown probability of the gap is defined as

$$S(u) = \frac{n_f}{N}$$

where $n_f$ represents the number of simulated failures recorded.

Computer simulation is particularly useful in assessing the validity of new and complex test procedures and in comparing the effectiveness of different test procedures. An example is shown in Figure 3.29 that compares different test procedures for low-probability withstand voltages that are briefly described in the following.

**3.7.3.1 Hancox Test Procedures (Hancox, 1958)**

Hancox test procedures consist of several sequences of increasing voltages to breakdown, with the initial test voltage increased successively by one step at each level until breakdown occurs at the application of the first impulse.

![Figure 3.29](image-url)
3.7.3.2 Fryxell (or $\alpha-\beta$) Test Procedures (Fryxell, 1966)

These test procedures consist of two sequences of test with increasing voltages with $n$ voltage applications at each level, with different values for $n$ in the two test sequences:

- In the alpha test series, the test starts from a voltage $U_0$ of low breakdown probability; the voltage is increased in steps of $\Delta u$ with $\alpha$-voltage applications at each level, until breakdown occurs at $U_\alpha$.
- The beta test series starts at the voltage $U_\alpha - 2\Delta u$; the voltage is increased in steps of $\Delta u$ with $\beta$-voltage applications at each level, until breakdown occurs at $U_\beta$. The critical withstand voltage is taken as the highest voltage reached in a beta test series of withstands.

3.7.3.3 Modified $\alpha-\beta$ Test Procedures (Rizk and Vincent, 1977)

These test procedures follow the same test procedures proposed by Fryxell. However, the critical withstand voltage is taken as the first complete withstand beta test series. The measured critical withstand voltage obtained in this procedure could be slightly lower than the one obtained in a Fryxell test.

Computer simulation, with its ability to make large number of simulations, is also useful in the estimation of the effective number of tests necessary to achieve the required confidence limits, as illustrated in Figure 3.30, which shows the improvement in the confidence limits with increasing number of tests performed. Its main advantage is in the economy of time and cost. The final assessment of a test procedure must be the actual test itself.

![Figure 3.30](image_url) Variations of the confidence limits with the number of tests. (From Trinh, N.G. and Vincent, C., *IEEE Trans. Power Appar. Syst.*, PAS-99, 711, 1980.)
3.8 Conclusions

From the present review of some applications of statistical analysis to high voltage engineering, it may be seen that statistical analysis is foremost a versatile working tool for analyzing test data. When properly applied, statistical analysis can help to

- Identify erroneous data and hence their separation from the main population
- Regroup data from various experiments to improve the total amount of data available and their statistical significance
- Extrapolate the observed behavior to untested conditions
- Provide a quantitative evaluation of the capacity of an insulation system to withstand specified operating conditions
- Sort hidden populations in a set of experimental data, giving a better insight of the phenomenon studied
- Optimize the number of tests and test samples, especially when destructive tests are to be avoided
- Evaluate and compare the statistical significance of test procedures
- Evaluate new test procedures

Its main drawback resides in the implicit assumption that the observed phenomenon remains the same over the whole range of conditions considered, even if they are beyond the effective limits covered by the experimental data, a condition not always met in practice.

REFERENCES


