66

STATISTICAL ANALYSIS OF DIELECTRIC TEST RESULTS

S. YAKOV
On behalf of
Working Group 01 / Task Force 02*
Of Study Committee 15 (Insulating Materials)

1991

cigré

STATISTICAL ANALYSIS
OF DIELECTRIC TEST RESULTS

S. YAKOV

on behalf of

Working Group 01/ Task Force 02*
of
Study Committee 15 (Insulating Materials)

1991

---

* **Members :** J. BINGGELI (Switzerland) - substituted later by P. BOSS (Switzerland), Y. KAKO (Japan), J. SAMAT (France) - substituted later by D. LACAZE (France), P. VUARCHEX (France) and S. YAKOV (Italy) - Convener

STATISTICAL ANALYSIS OF DIELECTRIC TEST RESULTS

Table of contents

1. Introduction

This paper is result of the work carried out by Task Force Nr 2 of CIGRE WG 15.01 "Fluid Impregnated Insulating Systems". The Task Force was set up with the purpose to study the determination of the dielectric withstand of fluid impregnated insulating systems.

Since the beginning of the work it became clear that the study of the dielectric strength of any kind of electrical insulation is based on dielectric tests on suitable samples and that both the design of the tests and the analysis of the results require the application of statistical methods. It was decided then to delimit the activity of the Task Force to the study of the statistical analysis of dielectric test results.

In the course of the work it was realized that the matter was already treated by CIGRE WG 33.03 "HV Testing and Measuring Techniques". A liaison was then established between the two CIGRE bodies as a result of which the papers prepared by each one are not redundant.

The WG 33.03 paper [1] is strictly related to the revision of Appendix A of IEC Publication 60-2 "High-Voltage Test Techniques" (1973). It contains a classification of HV test procedures and a short description of the analysis of test results by classical methods (momentum and regression estimation of parameters) and by the maximum likelihood method. The classical and the maximum likelihood methods are compared with the aim to show the advantages of the latter. Theoretical explanations are not given. Reference is made to the vast mathematical literature.

The nature of the present paper is different. It is intended as a guide for the analysis of the results of any kind of dielectric tests by means of statistical methods based on the principle of the maximum likelihood. The other classical methods, except the most simple graphical method, are not considered since the advantages of the likelihood approach have been shown in [1]. However, to facilitate the readers interested in the classical methods and to allow comparison with the likelihood approach, an extract of reference [1], regarding the classical methods, is reproduced here in Appendix 8.

The present paper has also tutorial character and therefore it contains more information, more detailed explanations and examples of practical application of the likelihood method to the most frequently used types of dielectric tests.

Moreover, in order to ease the understanding of the analysis and to reduce the possibilities of application and interpretation errors, some new terms are introduced, or more stringent definitions of existing ones are given in the paper (for ex. the term test-voltage and the classification of the tests in type P and type F in Sect. 2). This is in line with one of the aims of the paper that is to suggest a common language among the people interested in research on electrical insulation and in testing of models or complete equipment.

2. Dielectric tests

2.1 Test-voltages

The dielectric tests are characterized by the application of one or more test-voltages to the terminals of a test object according to a specified procedure.

In general the voltage applied during the test varies with time, $u=u(t)$, starting from an instant t_0 at a value U_0 and ending at an instant t_e at a value U_e. A **test-voltage** is defined completely by all the values of $u(t)$ in the interval t_0—t_e. Moreover it is necessary to ascertain that neither the results, nor their interpretation, are influenced by the events occurred before t_0 and after t_e. As concerns the effects on the insulation, it may be possible to simplify the definition of $U(t)$ giving only a few parameters, critical for the behavior of the insulation. Examples of such simplified definitions of the **test-voltage**[1] shape are the parameters which characterize the standard lightning and switching impulses specified in the IEC Publication 60-1 (1989-11) [3]. In other cases, for example in tests with ac or dc test-voltages, their shape should be completely defined as indicated below.

The most frequently used test-voltages are the following:

a) Impulse test-voltages

The impulse test-voltage, as specified in [3] is an aperiodic transient voltage which rises rapidly to a peak value and then falls more slowly to zero. It may have various shapes which are characterized by three parameters: front time (T_1), time to half value (T_2) and peak value (U_m). These parameters refer to "full impulses". There are also "chopped impulses" which are impulses suddenly interrupted by a disruptive discharge causing a rapid collapse of the voltage , practically to zero. Such impulses are characterized by two additional parameters: time to chopping (T_c) and steepness of the voltage collapse. Some of the impulses normally used in dielectric tests are:

- Standard lightning impulse (T_1=1.2μs, T_2=50μs)
- Standard lightning impulse chopped on the front ($T_c \leq 1.2$ μs)
- Standard lightning impulse chopped on the tail (T_c=2μs to 5μs)
- Standard switching impulse (250/2500μs)
- Special switching impulses (100/2500μs, 500/2500μs etc.)

---

[1] In this paper the term test-voltage (with the hyphen between the two words) indicates the concept defined in Section 2.1 above, otherwise the hyphen will be omitted. The term voltage stress, introduced in [2], is changed here into test-voltage to avoid confusion with voltage gradient (voltage/distance).

b) Constant test-voltages

A constant test-voltage is obtained with a dc voltage and consists of:

- a first part in which the voltage is increased up to the test level, usually linearly, starting from zero, or from a value sufficiently low to neglect its influence on the insulation;
- a second part in which the level is maintained constant until failure, or until a prefixed time limit t_l;
- a third part in which the voltage is decreased to zero gradually or abruptly, by opening the circuit.

The parameters that define the constant test-voltage are therefore: the rate of rise r of the voltage (in the first part), the voltage level U and the maximum duration t_l (in the second part).

Alternating voltages with peak values that vary according to the pattern of the three parts described above can be considered conventionally as constant test-voltages.

c) Stair test-voltages

A stair test-voltage is obtained with d.c. voltages and consists of parts in which the voltage is increased, usually linearly, followed by parts in which the voltage is kept constant as shown in Fig.1. The voltage is varied in this manner until the failure of the test object, or until the maximum voltage of the test equipment is reached without failure.

The stair test-voltage is defined by the following parameters:

U_1 - initial voltage level
ΔU - voltage step (voltage increment)
r - rate of rise of the voltage between two successive voltage levels
t_s - duration of each voltage level

Alternating voltages having peak values that vary according to the diagram in Fig.1 can be considered conventionally as stair test-voltages.

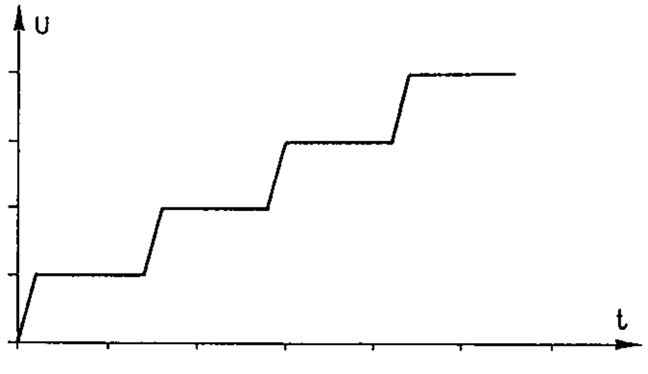


Fig.1 - Example of stair test-voltage

d) Ramp test-voltages

The ramp test-voltage is a continuously increasing voltage, usually starting from zero with constant rate of rise until the failure of the test object, or until the maximum voltage of the test equipment is reached without failure. The ramp test-voltage is defined therefore only by one parameter: the rate of rise r (expressed for ex. in kV/s). According to the value of r the ramp test-voltages can be of dc type or of impulse type. An example of this last type are the "linearly rising front chopped impulses indicated in [3].

Alternating voltages with continuously increasing peak values can be considered conventionally as ramp test-voltages.

e) Family of test-voltages

A family of test-voltages is a group of voltages of the same type in which one of the parameters defining the voltage takes various values. For example standard switching impulses with different peak values form one family of test-voltages.

Although the family of test-voltages, obtained from the variation of the peak value U_m for impulse test-voltages, or the voltage level U for constant test-voltages, is by far the most frequently used, there are also other families, based on the variation of parameters such as the rate of rise r of the ramp test-voltages, the front time T_1 of impulse test-voltages, the duration t_l of constant test-voltages etc.

f) Complex test-voltages

In some cases it is convenient to consider a particular sequence of test-voltages, pertaining to the same family, as one complex test-voltage. The parameters that define such complex test-voltage are:

- the type of the single test-voltages
- the values of the parameters that characterize the first single test-voltage applied and the rule for their variation in the sequence
- the time interval between the single successive test-voltages.

2.2 Insulation behavior under a test-voltage

The effects of a test-voltage on the insulation are quite complex and some of them not well known. However, from the practical point of view, it is expedient to define only two **possible outcomes of a test-voltage application: failure or withstand.**

The definition of failure depends on the phenomenon of interest. The failure can be defined as the total breakdown of the insulation, or as the appearance of partial discharges with apparent charge higher than a specified limit, or as the beginning of gas evolution in the insulation etc. etc..
The withstand can be defined as "no failure".

The physical processes that determine the failure are of a random nature. In consequence a statistical approach should be applied to the study of this phenomenon, considering two aspects:

- The simple fact that a failure occurred under a particular test-voltage. This is a random event characterized by its probability of occurrence P.
- The time t_f ($t_o \le t_f \le t_L$) at which the failure took place. This time is a random variable[2] to which are associated:
  - A probability distribution function F(t) giving, for every value t, the probability that $t_f \le t$.
  - A probability density function f(t) which is the derivative of the distribution function:

$$f(t) = \frac{dF(t)}{dt} \qquad (2.1)$$

Therefore dF(t)=f(t)dt is the infinitesimal probability that the random variable t_f take the value t ($t \le t_f \le t+dt$).

In many cases only one of the above aspects is taken into consideration.

The probability P of the event "failure" and anyone of the functions f(t) or F(t) characterize the response of the insulation to the applied test-voltage. Regarding the last two characteristics, for the sake of simplicity, in the following only the function F(t) will be considered. Obviously when F(t) is known it is possible to determine also f(t), if it exists.

It is important to emphasize here that any change in the shape of the test-voltages may change the characteristics P and F(t) which, therefore, depend on the applied test-voltage.

After a failure some types of insulation completely recover their insulating properties while others do not. They are named **self-restoring and non-self-restoring insulation** respectively.

In case of withstand it is possible that the insulation characteristics remain unchanged or they may change both negatively (deterioration) or positively (improvement or conditioning). It is very important to keep this in mind when a series of test-voltages is applied to an object during a dielectric test.

2.3 Procedures of test-voltage application and information that can be obtained

In the determination of the dielectric characteristics of an insulation one or more test-voltages are applied to it according to a specified procedure.

When the insulation characteristics do not change under the effect of the applied voltage, the test object can be submitted to the complete test procedure. Then, if the insulation is also self-restoring[3], several series of tests ending with failure can be repeated on the same object. If the insulation is non-self-restoring, the object must be changed with a nominally equal one after each failure.

When the insulation characteristics change under the effect of the applied voltage the object should not be submitted to the complete procedure and should be changed after each test-voltage application. However, there are cases when this rule is not applied, for example when the purpose of the tests is to investigate precisely how the insulation characteristics are affected by the applied test-voltages.

Generally the test procedures are based on the application of equal or different test-voltages. Let us examine the most common procedures and the information that is usually obtained from the corresponding test results.

A) Equal impulses

One frequently used procedure consists in the application to the test object of N impulses having the same shape and peak value U. The result of such experiment is a certain number of failures N_f and a certain number of withstands N_w=N-N_f. When, in the case of withstand, the characteristics of the insulation do not change under the effect of the applied test-voltages, the probability of failure of the test object relative to the application of one impulse can be estimated from the ratio:

$$P = \frac{N_f}{N} \qquad (2.2)$$

From the oscillographic record of each failure it is possible to measure the time to failure t and then, from the data obtained in this way, to estimate the corresponding probability distribution F(t).

B) Different impulses

The procedure described in A) can be repeated for different peak values of the impulses, maintaining constant all the other parameters. In other words the impulses are of a particular family of test-voltages.

If the insulation does not change its characteristics under the effect of the applied impulses, the results obtained permit the estimation of:

- the failure probability P, relative to the application of one impulse, as function of the peak value of the test-voltage: P=P(U);
- the probability distributions of the time to failure relative to different peak values: F(U,t).

---

[2] When the meaning of a term, related to probability and statistics, is not reported explicitly it is intended that the definitions given in the ISO Standard Nr 3534 (1977) [4] are applicable.

[3] Self-restoring refers to insulation tested in laboratory. The same insulation may not be self-restoring in actual service (for example because of the power in the arc ensuing a sparkover).

Three methods of test-voltage application are used in the procedures based on different impulses:

Multiple level method

A certain number of voltage levels is chosen and the same number of impulses N is applied at each level. The failure probability relative to the level U_i is estimated from the ratio:

$$P(U_i) = \frac{N_{fi}}{N} \qquad (2.3)$$

where N_{fi} is the number of failures observed at the level U_i.
The multiple level method is suitable for the estimation of the characteristic P(U) in the interval of voltage levels taken into consideration.

Step method

This method consists in the application of a family of impulses with peak values increased by steps until failure.
The failure probability corresponding to a given voltage level U_i (always on condition that the applied impulses do not change the insulation characteristics) can be estimated from the ratio:

$$P(U_i) = \frac{N_{fi}}{N_i} \qquad (2.4)$$

where N_i is the total number of impulses applied at the level U_i.
So, the characteristic P(U) can be obtained in the same way as in the case of the multiple level method. The difference between the two methods is in the value of N_i, which is constant in the multiple level method while in the step method it is not fixed in advance and is a result of the tests.

As effect of the step procedure N_i diminishes with the increase of the voltage level. In consequence the number of applied impulses is not distributed uniformly over the voltage range and the precision of the characteristic P(U) is better for the lower voltage levels. Therefore the method is to be used when we are interested in voltages having low failure probabilities.

Up and down method

This method is described in [3] and reported here in Appendix 1. Also in this case the number of applied impulses is not distributed uniformly over the test-voltage range. Most of the applied impulses result concentrated around a voltage level having a failure probability P, the value of which is selected in advance. So, the method is suitable for the determination of this particular characteristic: the voltage U_p to which corresponds a failure probability P.

C) Equal constant test-voltages

The outcome of the application of a constant test-voltage with level U, to the test object is a failure at the instant $t \leq t_L$ or no failure within the time limit t_L. In consequence the results of the application of N constant test-voltages having the same level U consist of N_f values of the time to failure (t_1, t_2, ..., t_{Nf}) and $N_w = N - N_f$ withstands i.e. N_w unknown values of t for which we know only that $t > t_L$. Tests of this kind, during which some withstands have occurred, are named **censored tests** and the data obtained **censored data.**

Arranging the times to failure in increasing order (i.e. $t_{j+1} > t_j$), the cumulative probability of failure $F(t_j)$ can be estimated from the ratio:

$$F(t_j) = \frac{1}{N}\sum_{i=1}^{j} N_i \qquad (2.5)$$

N_i being the number of failures occurred at the instant t_i. Since the time to failure can take any value in the interval 0—t_L the cumulative probability F(t) is a continuous function.

When the test results are censored it is possible to estimate the failure probability P, relative to the application of one constant test-voltage from expression (2.2) as in the case of equal impulses.

D) Different constant test-voltages

The procedure described in C) can be repeated at different levels of the test-voltage U. The insulation characteristics that can be estimated from the results are:

- the failure probability P, relative to the application of one constant test-voltage, as function of the voltage level: P=P(U);
- the probability distribution of the times to failure relative to each voltage level: F(U,t).

This procedure is equivalent to the multiple level method described in B). In fact all the procedures involving impulses are applicable also to the constant test-voltages. So, with the constant test-voltages it is possible to use the multiple level, the step and the up and down methods.

E) Equal ramp test-voltages

In the tests with ramp voltages the relation between the time t elapsed from the beginning of the test and the voltage u, applied to the test object, is well defined (usually u=rt). The outcome of the application of a ramp test-voltage, therefore, can be expressed either by the time to failure t or by the corresponding voltage u. So, the result of N ramp tests will be a series of values of the above quantities.

If the maximum voltage U_L of the test equipment is not sufficiently high to cause always the failure of the test object, the result of the application of N ramps will consist of N_f values of the failure voltage u_1, u_2, ... ,u_{Nf} and of $N_w = N - N_f$ unknown values of u about which we know only that they are higher than the limiting voltage U_L. Here again we have the case of censored data.

From the test results it is possible to estimate the failure probability distributions F(u) or F(t) in the same way as explained in C).

F) Equal stair test-voltages

If the results of the tests with equal stair voltages are expressed as a series of values of the time to failure, the considerations made about the ramp test-voltages are directly applicable to the stair test-voltages.

However it should be noted that, in the stair tests, to any value of t corresponds a well defined value of U but not vice versa (see Fig.1). In other words the relation between voltage and time is not "bi--univocal" as in the ramp test. Therefore some information is lost when the results of the stair tests are expressed only by the voltage level at which the failure occurred. In spite of that, it is more usual to present the results as series of failure voltages and, when a failure occurs during the variation of the voltage from the level U_j to the level U_{j+1}, it is generally accepted to report conventionally as failure voltage one of the two levels, for example U_{j+1}. In this way the failure voltage becomes a discrete variable and, in order to keep this in mind, the corresponding cumulative probability function will be indicated with $F(U_j)$.

G) Equal complex test-voltages

The most frequently used complex test-voltage consists of a series of impulses, or constant voltages, with increasing amplitudes.

The complex test-voltage, consisting of series of constant voltages with increasing amplitudes is similar to the stair test-voltage. So, it can be named **complex stair**. The difference between the two test-voltages can be seen from Fig.2 which shows the complex stair in which there are intervals with zero voltage between the single constant voltages while in the stair test-voltage, illustrated in Fig.1, such intervals do not exist. The term complex stair can be applied also when the single voltages in the series are impulses.

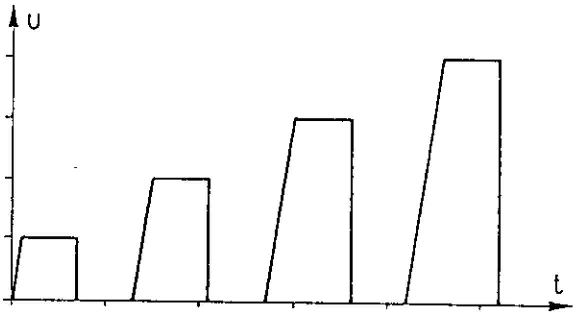


Fig.2 - Example of voltage varied according to the step method

The characteristic determined with the application of equal complex stairs is the cumulative probability $F(U_j)$, which is a discrete function, as explained in F) for the simple stair procedure.

N.B. Attention should be paid to avoid confusion between the step method and the test procedure with equal complex stairs. In fact, in both cases impulses, or constant voltages, with increasing amplitudes are applied to the test object. However, the same set of experimental data can be viewed in two ways:

- **as data obtained applying different test-voltages that are simple** and, in such case, the characteristic P(U) is estimated,

or

- **as data obtained applying equal test-voltages that are complex** and, in such case, the characteristic $F(U_j)$ is estimated. The last estimation can be made also when the insulation characteristics change under the effect of the applied voltages.

Results of this kind are analyzed in Appendix 7 - Example Nr 7.

- . -

The test procedures discussed above and the information that can be obtained are summarized in Table 1.

In many cases not all the information that can be obtained from a test is neces-

TABLE 1 INFORMATION OBTAINED WITH THE DIFFERENT TEST PROCEDURES

TEST-VOLTAGES	TEST-PROCEDURES			
	EQUAL	DIFFERENT TEST-VOLTAGES		
	TEST-VOLTAGES	MULT.LEV.	UP & DOWN	STEP
Impulse	P, F(t)	P(U), F(U,t)	U_p	P(U), F(U,t)
Constant	F(t)	P(U)	U_p	P(U), F(U,t)
Ramp	F(u), F(t)	-	-	-
Stair	$F(U_j)$, F(t)	-	-	-
Complex stair	$F(U_j)$	-	-	-

sary. For example, in impulse tests the time to failure is often of no interest and what we are looking for is the peak value of the impulse which causes the failure i.e. the characteristic P(U). On the contrary, in the tests with constant test-voltage we are usually more interested in the characteristic F(t). So, according to the characteristics we want to estimate from the experimental results, i.e. according to the purpose of the tests, the last can be classified in two groups: tests of the type P and tests of the type F. It is expedient to emphasize here the differences between the two tests and the two insulation characteristics.

In **tests of type P** the information is obtained applying a series of equal test-voltages to the object. So, these tests are named also **constant-stress tests. The insulation characteristic of the type P** represents the relationship between the variable and the failure probability P. For example, in the case of impulse tests, P(U) is the probability of failure when one impulse with crest value U is applied.

In the **tests of type F** the variable X (equal to U or t) is increasing in the course of the test and therefore these tests are named also **progressive-stress tests.** **The insulation characteristic of the type F** is a distribution function i.e. a function F(X) giving, for every value X, the probability that, in a test, the random variable be less then or equal to X. By definition F(X) is always increasing while this statement is not valid "a priori" for P(U).

The information obtained from dielectric tests, performed according to the procedures described here, can be used for various purposes such as insulation design and coordination, comparison, acceptance or rejection of equipment etc.. It is important to select the procedures that suit better the purpose we have in mind. However, the detailed examination of the various purposes that the dielectric tests may have and the choice of the most suitable test procedure is not within the scope of this paper.

The principal aim of the paper is to present a method for the estimation of the insulation characteristics of the type P and of the type F by means of statistical analysis of experimental results.

An estimation of the characteristics of the type P and F can be done graphically. Such estimation is very simple but has some disadvantages which can be avoided by the statistical analysis at the cost of a greater complexity.

For the better illustration of the statistical analysis it is expedient to examine first the graphical approach and its problems.

3. Graphical estimation of insulation characteristics

3.1 Characteristics of the type P

Let us consider an experiment consisting of a series of tests with impulses applied according to the "multiple level method" (see sect. 2.3 B). A certain number of voltage levels (impulse peak values) are selected and N impulses of each level are applied to the test object. Suppose that the insulation of the object does not change its characteristics under the effect of the applied impulses (except in case of failure). The tests, therefore, are mutually independent.

An example of results obtained in this way is given in Table 2. It refers to tests with special switching impulses 400/4000 μs of positive polarity on a 2 m rod-plane gap in air. The results relative to the level U_i consist of a number of failures N_{fi} and a number of withstands $N_{wi}=N-N_{fi}$. The relative frequency of failure $P_{ei}=N_{fi}/N$ is an estimation of the failure probability P_i, associated with the application of one impulse having peak value U_i. In Fig.3 the estimated values P_e of the failure probability are indicated with • and plotted versus the peak value U of the impulses.

One of the problems to be considered in the estimation of the insulation characteristics is related to the "uncertainty" of the experimental results. In fact, if the experiment is repeated the results will not necessarily be identical with those of the first experiment because of the random character of the failure event. The failure frequency N_f is a random variable and since the tests are independent the corresponding probability distribution is that of **Bernoulli** (or the **binomial distribution**). So, repeating the experiment several times, we shall obtain a set of different estima-

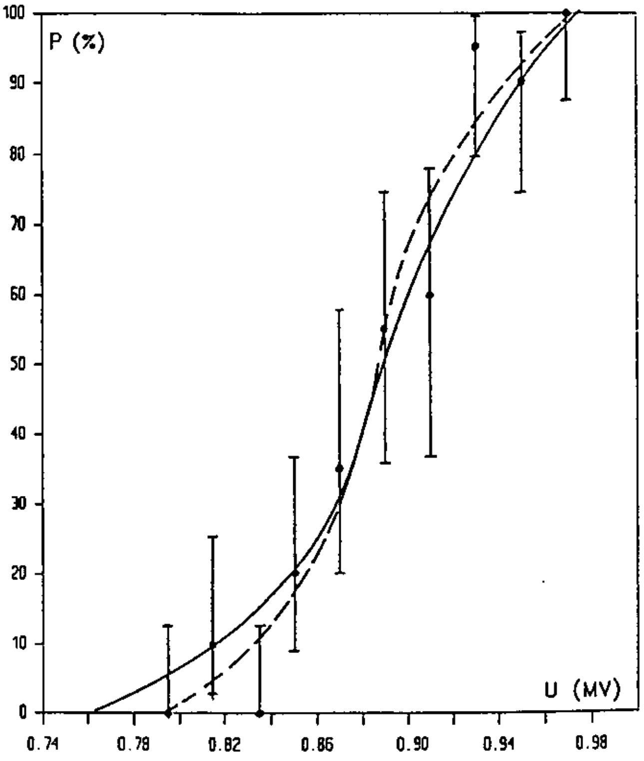


Fig.3 - Failure probability of type P

tions P_{ei} of the failure probability P_i. The "true" value P_i can be obtained only from an experiment consisting of a very high (theoretically infinite) number of tests. Obviously, this is not possible and, therefore, it is important to have an idea about the amount of "uncertainty" in the estimation of the failure probabilities.

Making use of the experimental results and the binomial distribution we can construct an interval about the estimated value P_e and state, with a specified confidence C, that the "true" value P of the failure probability lies in the interval. Such interval is called **100·C% confidence interval.** For a 90% confidence interval, the statement will be correct in about 90% of the times in the sense that, if we make numerous replications of the experiment and compute each time the corresponding confidence interval, in about 90 cases out of 100 the confidence intervals will contain P.

The end points of the confidence interval are called **confidence limits.** These limits depend on the number of failures N_f, on the total number of tests N and on the confidence C[4]. They can be obtained from tables for "two sided confidence limits for a proportion" available in many statistical manuals. An example is Table 21 taken from [5] and reproduced here in Appendix 2.

In Fig.3 the values of P_e are shown together with their confidence intervals.

TABLE 2 EXAMPLE OF TYPE P TEST

Test object:	Rod-plane 2 m air gap
Test-voltage:	SI + 400/4000 μs
Test conditions:	Dry tests; N=20 for all voltage levels

U (kV)	N_f	P_e (%)	90% CONF. LIM. (%)	
			LOWER	UPPER
795	0	0	0	12.6
815	2	10	2.7	25.5
835	0	0	0	12.6
850	4	20	9.0	36.7
870	7	35	20.1	57.8
890	11	55	35.8	74.5
910	12	60	36.7	77.9
930	19	95	79.7	99.5
950	18	90	74.5	97.3
970	20	100	87.4	100.00

Another problem is presented by the need of information about the insulation behavior when impulses having the same shape but peak values different from those of the experiment are applied. Such information can be obtained from a great number of experiments but, normally, this way is not practical due to the high costs involved. An alternative approach to the problem is based on the plotting of an "interpolating curve" between the experimental points. For the purpose a certain number of assumptions about the function P(U) are made. The experience has shown that it is acceptable to consider this function continuous, smooth and, practically in all cases, increasing with U. Then it is possible to plot "by eye" a curve having these characteristics and satisfying the condition to lay inside the majority of the confidence intervals of the experimental points. This, however, is not always easy.

In Fig.3 are shown two "possible" curves and it is difficult to decide which one is better. Obviously the graphical procedure is rather subjective because there is an infinite number of curves that satisfy the requirements specified above and the selection between them is arbitrary.

3.2 Characteristics of the type F

For the illustration of the graphical procedure suitable for the estimation of insulation characteristics of type F, let us consider tests with ramp voltages.

In Table 3 is given an example of results obtained from tests on 50 mm point-to-sphere gap in transformer oil with 50Hz ramp voltage having rate of rise r=1kV/s. In these tests the appearance of a partial discharge pulse with apparent charge exceeding 100 pC was defined as failure. Since the oil insulation of the gap, tested in these conditions is self-restoring, 20 tests have been performed on the same object.

TABLE 3 EXAMPLE OF TYPE F TEST

Test object:	50 mm point-sphere gap in oil
Test-voltage:	50Hz ramp r=1 kV/s
Failure criterion:	PD ≥ 100 pC

i -	U_i (kV)	N_i -	ΣN_i -	$F_e(U_i)$ (%)
1	20.5	1	1	5
2	21.7	1	2	10
3	22.2	1	3	15
4	22.3	1	4	20
5	22.4	1	5	25
6	22.8	4	9	45
7	23.3	1	10	50
8	23.4	2	12	60
9	23.5	2	14	70
10	23.8	1	15	75
11	23.9	1	16	80
12	24.3	1	17	85
13	24.6	1	18	90
14	25.0	1	19	95
15	25.2	1	20	100.

---

[4] C is named confidence coefficient in the statistical literature.

In Table 3 are given the failure voltages, arranged in increasing order, their frequency of occurrence N_i, the cumulative frequency ΣN_i and the estimated values $F_e=\Sigma N_i/20$ of the probabilities $F(U_i)$. In Fig.4 is shown the step diagram corresponding to the values of $F_e(U_i)$. This diagram is called **empirical distribution function** and represents the estimation of the function F(U). If the number of tests is increased the number of steps in the diagram will increase and their height will decrease. For an infinite number of tests the step diagram will become a smooth curve corresponding to F(U). So, if we prefer, we can substitute the empirical distribution with a smooth curve placed somehow within the steps and consider this last curve as estimation of F(U). Here again we are faced with plotting difficulties because the shape of the curve and its position in respect of the step diagram are not completely defined.

Since the estimation of F(U) is based on a finite number of tests we have to keep in mind that it is uncertain. In fact, repeated tests usually give different results and, therefore, different empirical distributions. Thus, also in this case, it is important to have an idea about the amount of "uncertainty" in the estimation of the function F(U). For the purpose a **"confidence band"** can be formed around the empirical distribution function, constructed in such a way as to obtain a 100·C% probability that the unknown distribution function F(U) lies entirely within its bounds.

The bounds of the confidence band can be determined as follows:

upper bound = $F_e(U_i) + w(N,C)$
lower bound = $F_e(U_i) - w(N,C)$

where w(N,C) is the Kolmogorov test statistic which depends on the number of results N and on the confidence coefficient C. The values of w(N,C) can be obtained from Table 14 taken from [7] and reproduced here in Appendix 3. Obviously the upper bound can not exceed 1 and the lower bound can not be extended below 0.

Let us determine the 90% confidence bounds of $F_e(U)$ relative to the data of Table 3. From the table of Appendix 3 it can be seen that for N=20 and C=0.9 the value of the Kolmogorov test statistic for two sided test is[5]: w(20,0.9) = 0.265 (26.5%). The confidence bounds corresponding to this value of w are plotted in Fig.4.

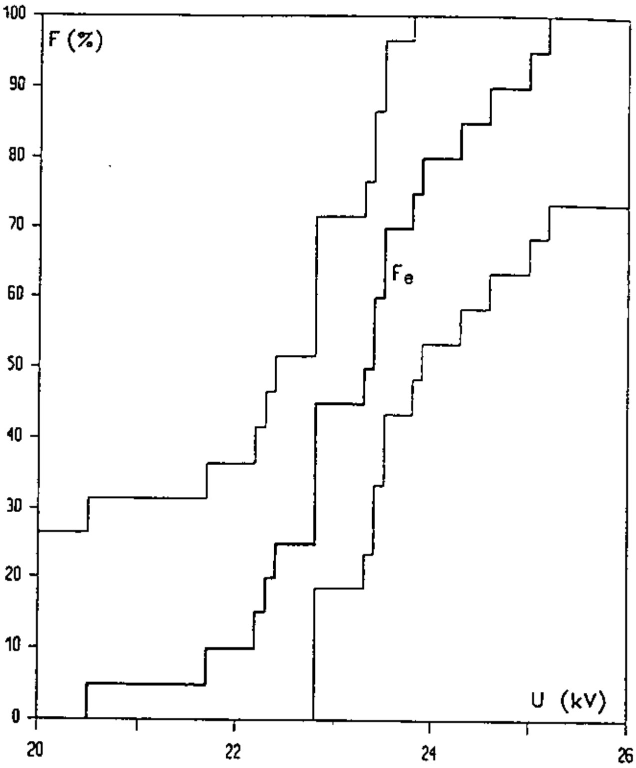


Fig.4 - Failure probability of type F

3.3 Some comments about the graphical estimation

The graphical estimation of insulation characteristics of the type P and F has the advantage to be very simple. The data treatment consists only in the determination of the relative frequencies N_{fi}/N_i or $\Sigma N_i/N$. No other calculations are needed. The information for the determination of the confidence limits and bounds is obtained from tables available in the statistical literature. Only one hypothesis is made when the estimated curves of the failure probability are plotted, namely the hypothesis that the functions P(U) and F(U) are continuous and smooth. F(U) is increasing with U by definition while for P(U) such an hypothesis is known to be reasonable from experience.

The main disadvantage of the graphical estimation is the subjective character of the curve plotting. In fact an infinite number of curves can be plotted trough the experimental points, in the case of a P characteristic, or trough the step diagram of the empirical distribution, in the case of an F characteristic. The selection of both the shape and the position of the curve is arbitrary.

A more objective method of curve plotting is given by the **"statistical interpolation"** of the experimental results which consists in the following steps:

- Selection of the type of interpolating function usually specified by a mathematical expression. The last depends on the variable and on a certain number of parameters the values of which are not known.
- Selection of a criterion for the **"best fitting"** of the interpolating function to the test data.
- Determination of the values of the parameters of the interpolating function from the test data in manner to satisfy the selected criterion for curve fitting.

These steps are examined in the next sections.

---

[5] In Appendix 2 the symbols used for N and C are n and p respectively.

4. Interpolating functions and probability papers

The best functions to use for the interpolation of the experimental results are the ones that reflect the physical laws which govern the failure of the insulation. Unfortunately, in most of the practical cases, these laws are not known. Therefore, any mathematical expression, to which corresponds a curve that fits sufficiently well the experimental data, can be used as interpolating function.

At first, one of the most important probability distribution functions used in statistics i.e. the Gaussian, or Normal, distribution was adopted. For several years this distribution was employed in the analysis of the results of dielectric tests and the engineers involved in dielectric problems got used to the parameters that define it (i.e. the mean value and the standard deviation).

Latter it was found that other distributions, namely those of Gumbel and Weibull may, in many cases, fit better the experimental data. The mathematical expressions of these two distributions and that of the Normal one are reported in Table 4. In this table the independent variable is indicated with X. So, this symbol may represent voltage or time according to the type of experimental results which are analyzed.

The probability density φ and the probability distribution Φ are given as functions of the so called "**standardized variate**" y which, on its turn, is a function of X. The functions y(X) are also shown in Table 4. It can be seen that they depend on two or three parameters.

For the Gumbel and the Weibull distributions in the table are given both the "classic" expressions, as presented in the relevant literature, and the "modified" ones. The last are obtained from the classic expressions with a change of parameters in manner to characterize all the three distributions with the more familiar parameters of the Normal distribution. The derivation of the modified expressions is shown in Appendix 4.

In the literature the parameters of the Normal distribution are defined as mean (or expectation), denoted usually with μ, and standard deviation, denoted with σ.

When characteristics of the type F are considered, the values of μ and σ can be estimated from the test results by means of the expressions:

$$\mu = \frac{\sum X_i}{N} \qquad \sigma^2 = \frac{\sum (X_i - \mu)^2}{N - 1} \qquad (4.1)$$

However, these expressions can not be applied to censored data and when characteristics of the type P are considered. In fact the voltage levels in type P tests are selected by the operator and their mean value and standard deviation have nothing to do with the function P(U).

Since the functions $\Phi(X)$ of Table 4 are used for the interpolation of results of both type P and type F tests, it is important to avoid possibilities of misinterpretation and errors. For this reason it is preferred here to avoid the use of the mean and the standard deviation as parameters and to adopt in their place the following quantities:

X_{50} - the value of X to which corresponds a probability of 50%
Z - the difference between X_{50} and X_z, the last being the value of X to which corresponds a probability of 15.87%.

These parameters have been chosen because one of the characteristics of the Normal distribution is that X_{50} is equal numerically to the mean, to the median and to the mode[6]: $X_{50} = \mu = X_{md} = X_{mo}$. For the Gumbel and the Weibull distributions X_{50} coincides with the median ($X_{50}=X_{md}$) but the mean, the median and the mode do not coincide. The peculiar value of 15.87% has been chosen for $\Phi(X_z)$ in order to obtain $Z = \sigma$ in the case of Normal distribution. In this way the parameters X_{50} and Z, in spite of their different meaning, remain familiar to those engineers who are accustomed with the use of the Normal distribution and, what is more important, they are directly comparable to all the values of μ and σ available in earlier publications.

From Table 4 it can be seen that the Weibull distribution depends on three parameters. In the classic expression these parameters are:

α - exponent, called also shape parameter because it defines the shape of the curves $\varphi(X)$ and $\Phi(X)$
X_s - scale parameter; $X_s = X_{mo} - X_o$
X_o - location or threshold parameter; to $X \le X_o$ corresponds $\Phi(X_o)=0$.

The presence of the threshold parameter X_o makes the Weibull distribution more attractive to the engineers because it corresponds to the intuitive idea of a withstand voltage. With the other two distributions the withstand voltage should be defined as a voltage to which corresponds a certain, sufficiently small failure probability. The choice of how small should be this probability remains arbitrary.

On the other hand, the presence of three parameters makes the use of the Weibull distribution more difficult. In order to avoid this inconvenience, very often it is assumed that X_o is equal to zero. This simplifies the mathematics but one of the attractive characteristics of the Weibull distribution is lost.

In the modified expression of the Weibull distribution the threshold parameter is defined as follows:

$$X_o = X_{50} - K_o Z \qquad (4.2)$$

---

[6] The mode X_{mo} is the value of X for which the probability density function $\varphi(X)$ is maximum.

TABLE 4 INTERPOLATING FUNCTIONS

FUNCTION	GAUSS (NORMAL)	GUMBEL		WEIBULL	
		CLASSIC	MODIFIED	CLASSIC	MODIFIED
PROBABILITY DENSITY $\varphi(X)=\varphi[y(X)]$	$0.40y'\exp(y^2/2)$	$y'\exp[y-\exp(y)]$	$0.69y'\exp(y)0.5^{\exp(y)}$	$y'\exp(-y)$	$0.69y'(0.5)^y$
PROBABILITY DISTRIBUTION $\Phi(X)=\Phi[y(X)]$	$0.40\int_{-\infty}^{y}\exp(t^2/2)dt$	$1 - \exp[-\exp(y)]$	$1 - 0.5^{\exp(y)}$	$1 - \exp(-y)$	$1 - 0.5^y$
STANDARDIZED VARIATE $y(X)$	$(X-X_{50})/Z$	$\alpha(X-X_{mo})$	$1.39\ (X-X_{50})/Z$	$[(X-X_o)/X_s]^\alpha$	$[1+(X-X_{50})/K_oZ]^\alpha$ $\alpha = 1.39/\ln[K_o/(K_o-1)]$
$y' = dy/dX$	$1/Z$	α	$1.39/Z$	$(\alpha/X_s)[(X-X_o)/X_s]^{\alpha-1}$	$(\alpha/K_oZ)[1+(X-X_{50})/K_oZ]^{\alpha-1}$
PARAMETERS	X_{50}, Z	α, X_{mo}	X_{50}, Z	α, X_s, X_o	X_{50}, Z, K_o

where K_o is a new parameter which indicates the "distance" between the median and the threshold values of X in terms of Z. Having defined X_o in this way, the following relation is obtained (see Appendix 4) between the exponent α and the parameter K_o:

$$\alpha = \frac{1.39}{\ln\left(\frac{K_o}{K_o - 1}\right)} \tag{4.3}$$

Therefore, the functions φ and Φ still depend on three parameters, namely on X_{50}, Z and K_o. One advantage of this change is the possibility to give reasonable values to K_o (for example in the range 3 to 5) and to reduce in this way the number of the parameters to two and still having $U_o \neq 0$.

For each interpolating function of Table 4 it is possible to construct a probability paper which is an useful tool for testing the fit between the experimental data and the selected function.

The probability paper is a rectangular grid where the variable X is plotted on one axis (usually the horizontal) and the probability on the other axis (usually the vertical). The scales of the two axes are chosen in such a way as to represent the selected interpolating function $\Phi(y)$ by a straight line when plotted against X. The construction of the probability papers is explained in Appendix 5.

The primary use of the probability paper is to assess the adequacy of the selected interpolating function to represent the test data. For the purpose the test results are first plotted on the corresponding probability paper. In fact the choice of a probability paper is identical with the choice of an interpolating function. If the results follow a fairly straight line, then it may be reasonable to assume that the selected function is adequate.

The probability paper is an improvement of the graphical method because it eliminates the plotting difficulties related to the shape of the interpolating curve. This improvement, however, is obtained at the cost of an hypothesis about the type of the interpolating function.

5. Criteria for fitting a function to test data

The problem of fitting an interpolating curve through the experimental data is not completely resolved by the transformation of the curve in straight line on the probability paper. The question is how to draw such a line through experimental data that generally do not fall exactly on a straight line. An infinite number of lines can be plotted through the data and the problem is to identify the "best" one. For the purpose it is necessary to define a criterion for the "best fitting". The most used one is that of the "**least squares**" but it is not applicable to all kinds of results that can be obtained from dielectric tests and, in particular, to censored data [7]. For this and many other reasons, it is recommended in this paper to use the criterion of the "**maximum likelihood**". This criterion is based on the "principle of the maximum likelihood" which presumes that the function which gives the "best fitting" to the experimental data is the one to which corresponds the maximum probability of obtaining these data. The maximum likelihood method for curve fitting can be summarized as follows:

- An hypothesis is made about the type of interpolating function $\Phi(X)$.
- Assuming that $\Phi(X)$ is true it is possible to determine the probability P_r of obtaining the results at hand.
- The mathematical expression of P_r depends on the results, which are known and on the parameters of $\Phi(X)$, the values of which are not known and have to be determined.
- An infinite number of hypotheses can be done about the values of the parameters of $\Phi(X)$ and for each one it is possible to calculate P_r.
- The likelihood of an hypothesis is defined as a quantity L proportional to the probability of obtaining the results, the constant of proportionality K being arbitrary: $L = KP_r$. The likelihood L, being a function of the parameters, is known also with the name **likelihood function**.
- In the comparison of two rival hypotheses to be preferred is the one with the higher likelihood. Therefore the "best values" of the parameters of $\Phi(X)$ are the ones for which the likelihood function L attains its maximum.

The maximum likelihood method is advantageous because it gives a very general solution to the curve fitting problem. Any type of interpolating function can be chosen and it is not necessary to "linearize" it for the determination of the parameters. The use of suitable probability papers, therefore, is optional. The method is applicable to all kinds of test results including censored data, values of P_e equal to 0 or 100% etc. Moreover, the maximum likelihood principle can be applied in the choice between different interpolating functions, in the determination of confidence bands and limits etc.. It is the basis of a general approach to the solution of the problems related to all the steps of the statistical analysis of dielectric test data. The illustration of this approach is the subject of the next sections.

6. Likelihood functions

Let us select an interpolating function $\Phi(X;A,B)$[7] and assume that it is the true representation of the failure probability P(U) or F(X).

---

[7] Although only two parameters A and B are indicated, the considerations that follow are valid for any number of parameters.

The likelihood L of an hypothesis H about the values of the parameters (H: $A=A_h$, $B=B_h$), given the set of results R and the type of probability function Φ, is defined as a quantity proportional to the probability of obtaining the results P_r, the constant of proportionality K being arbitrary:

$$L(R;A_h,B_h) = KP_r(R;A_h,B_h)$$

Similar expressions can be written for any other hypothesis about the values of the parameters so that in general we have:

$$L(R;A,B) = KP_r(R;A,B) \qquad (6.1)$$

As indicated in [8], the arbitrary constant K enables us to use the same definition of likelihood for discrete and continuous variables alike, and is no impediment to its use which invariably involves comparisons of likelihoods. In fact K is a constant in any one application, involving many different hypotheses but the same data and probability model. It is, of course, not necessarily the same constant in another application. This too is no hindrance because there is no need to make an absolute comparison of different hypotheses on different data.

The difference between likelihood and probability is explained in [8] as follows:

"Whereas with probability, the results R are variable and the hypothesis H (i.e. the values of A and B) is constant, with likelihood H is the variable for constant R. ...

The distinction between probability and likelihood is vital to an understanding of the role played by each in inductive inference. When considered as a function of R, the probability of obtaining the results $P_r(R;A,B)$ defines a statistical distribution, either discrete or continuous. As such, if we sum or integrate (as appropriate) over all possible results R we will obtain unity, by one of the axioms of the probability. Likelihood, on the other hand, is predicted on fixed data R, and for varying hypotheses may be regarded as a function of the hypotheses (or the parameters). But this function in no sense gives rise to a statistical distribution, and there is nothing in its definition which implies that if summed over all possible hypotheses (even if these could be contemplated) or integrated over all possible parameter values the result will be anything in particular."

In dielectric tests the probability of obtaining the results P_r and the likelihood L depend on the procedure of the test-voltage application and therefore their expressions must be determined for each type of test.

6.1 Type P tests (constant-stress tests)

In type P tests the same test-voltage is applied many times to the test object. Only two outcomes are considered for each application: failure or withstand. If the probability of occurrence of one event E_2 is independent of whether or not another event E_1 has occurred, the event E_2 is said to be statistically independent of the event E_1. For such events is applicable the theorem of the product: **the probability of the simultaneous occurrence of a number of statistically independent events is equal to the product of the probabilities of each event occurring separately:**

$$P(E_1E_2E_3\ldots) = P(E_1)P(E_2)P(E_3)\ldots$$

Let P denote the probability of failure. Then the probability of withstand will be Q=1-P. In an experiment consisting of N tests, the probability of obtaining N_f failures and $N_w=N-N_f$ withstands, in a particular order, according to the above theorem is:

$$P_{ro} = P^{N_f}Q^{N_w} \qquad (6.2)$$

Usually we are not interested in the order in which the failures occur. The number of orders in which N_f failures and N_w withstands can occur is:

$$C^N{}_{Nf} = \frac{N!}{N_f!N_w!}$$

These orders are mutually exclusive events, so the probability of exactly N_f failures (in N independent tests), irrespective of order, is the sum of the probabilities for all the $C^N{}_{Nf}$ orders:

$$P_r = \frac{N!}{N_f!N_w!}P^{N_f}Q^{N_w} \qquad (6.3)$$

The expression (6.3) represents the probability of obtaining the results in type P tests performed with equal test-voltages. In the determination of the corresponding likelihood function L it is expedient to choose $K=(N_f!N_w!)/N!$ in order to obtain the simple expression:

$$L = P^{N_f}Q^{N_w} \qquad (6.4)$$

As indicated in Section 2.3 B) and D), tests with equal test-voltages can be repeated at different levels. Let us consider an experiment consisting of tests performed at k different voltage levels. For each level U_i the probability of obtaining the results will be P_{ri} as given by (6.3). Since the tests at the various levels are statistically independent events, the theorem of the product applies and the probability of obtaining the whole set of results relative to all levels is:

$$P_r = \prod_{i=1}^{k} P_{ri} = \prod_{i=1}^{k} \frac{N_i!}{N_{fi}!N_{wi}!} P^{N_{fi}} Q^{N_{wi}} \qquad (6.5)$$

Since we have assumed that $P(U)=\Phi(X;A,B)$, it can be seen that the above expression depends on the parameters A,B and on the results N_f, N_w.

With a suitable choice for K we obtain for the likelihood function relative to "multiple level tests":

$$L = \prod_{i=1}^{k} L_i = \prod_{i=1}^{k} P_i^{N_{fi}} Q_i^{N_{wi}} \qquad (6.6)$$

6.2 Type F tests (progressive-stress tests)

In type F tests the variable X (voltage or time) is increased until failure. The result of an experiment consisting of N stress applications is a set of values of X: X_1 with frequency N_1, X_2 with frequency N_2, ... , X_m with frequency N_m. Obviously $m \le N$ and $N_1+N_2+\ldots+N_m=N$.

We have assumed that $F(X)=\Phi(X;A,B)$ and therefore for the probability of obtaining the result X_i in a given test we can write:

$$P(X=X_i) = \varphi(X_i;A,B)\,dX \qquad (6.7)$$

where $\varphi=d\Phi/dX$. The probability of obtaining the whole set of results (X_i,N_i), irrespective of the order in which they occurred is then:

$$P_r = \frac{N!\,(dX)^N}{N_1!N_2!\ldots N_m!} \prod_{i=1}^{m} \varphi(X_i;A,B) \qquad (6.8)$$

If we choose $K=[(N_1!N_2!\ldots N_m!)/N!]/(dX)^N$, we obtain for the likelihood:

$$L = \prod_{i=1}^{m} \varphi(X_i;A,B)^{N_i} \qquad (6.9)$$

This is the classic expression of the likelihood function given in the relevant literature. However, it is not applicable to the discrete probability functions $F(U_j)$ relative to the stair and complex stair test-voltages described in Section 2.3. In fact the derivative $\varphi(X_i;A,B)$ has no sense for discrete functions. Therefore an alternative expression for the probability of obtaining the results, valid both for continuous and discrete functions, is determined as follows.

In the case of continuous functions F(X), the X-axis is divided in a certain number of intervals which cover the whole range X_{min}—X_{max} of possible values of the variable (usually $X_{min}=0$ and $X_{max}=\infty$). The choice of the number and the size of these intervals is arbitrary but it is convenient that each interval contains at least one result[8]. The maximum number of intervals, determined according to this criterion, is obviously equal to m (the number of different values of X). Let the values of X be arranged in increasing order, indicating with X_1 the minimum and with X_m the maximum value, that is: $X_1 \le X_2 \le X_3 \le \ldots \le X_m$. The lower bound of the first interval is X_{min} (0). For the upper bound we can choose any point between X_1 and X_2 say the point $x_1=(X_1+X_2)/2$. Then the lower limit of the second interval will be x_1 and the upper $x_2=(X_2+X_3)/2$. The limits of the third interval will be x_2 and x_3 etc. as shown in Fig.5. Obviously the upper limit of the last interval is $X_{max}(\infty)$.

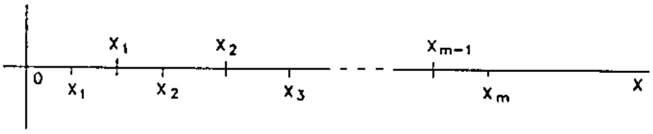


Fig.5 - Determination of the m intervals on the X-axis for type F tests

The probability of obtaining N_i results in the interval "i" is:

$$[\Phi(x_i;A,B) - \Phi(x_{i-1};A,B)]^{N_i} = \Delta\Phi_i(A,B)^{N_i}$$

It should be kept in mind that for i=1 in the above expression we have $x_0=X_{min}$ and $\Phi(x_0)=0$.

The probability to obtain the whole set of results, irrespective of the order in which they occurred, is:

$$P_r = \frac{N!}{N_1!N_2!\ldots N_m!} \prod_{i=1}^{m} \Delta\Phi_i(A,B)^{N_i} \qquad (6.10)$$

Using the same criterion as before in the choice of the value of K, we obtain for the likelihood function:

$$L = \prod_{i=1}^{m} \Delta\Phi_i(A,B)^{N_i} \qquad (6.11)$$

It should be noted that $\Phi(x_m=\infty)=1$ and therefore $\Delta\Phi_m=1-\Phi(x_{m-1})$.

---

[8] This criterion is used mainly for small number of results, say $5 \le N \le 20$. In general the question which is the "best minimum number of results" to include in an interval needs further studies.

The expression (6.11) remains valid also for the discrete functions $F(U_j)$. In such case the U axis is divided again in intervals, each containing at least one result, but the upper limits of these intervals must coincide with the test-voltage level it contains, that is (with reference to Fig.5):

$$x_j = X_j = U_j \qquad (6.12)$$

The reason for this choice is that the probability functions of Table 4, which are continuous, are used also to represent the discrete functions $F(U_j)$. When we do this we make the hypothesis about the existence of a discrete function which, for each value $X_j=U_j$ of the test-voltage, gives the same cumulative probability as the continuous functions $\Phi(X)$ of Table 4.

In this way, using expression (6.11) for the likelihood, the calculations for ramp, constant voltage, stair and complex stair tests are the same.

More comments about the two likelihood expressions (6.9) and (6.11) are made in Section 12.

6.3 Censored tests

As mentioned in Section 2.3, there are type F tests in which the variable X may have an upper limit X_L and the test is stopped when this limit is reached even if no failure has occurred. Examples of such "censored tests" are those with ramp and stair test-voltages when the maximum voltage of the test equipment is not sufficient to cause always the failure of the test object. A more frequent situation is represented by the tests with constant voltages of limited duration.

In all the cases indicated above the experimental data consist of a certain number of failures N_f at different values of X (X_1 with frequency N_1, X_2 with frequency N_2, ... , X_m with frequency N_m) and a certain number of withstands about which the only thing we know is that $X>X_L$ (censored data). Obviously $N_f=N_1+N_2+...+N_m$ and $N_w=N-N_f$ where N is the total number of tests.

The probability of obtaining the N_f failures at the levels X_1, X_2, ... , X_m is proportional to the likelihood function given by expression (6.9) and the probability of obtaining N_w results in the interval from X_L to ∞ is proportional to $[1-\Phi(X_L)]^{N_w}$. The probability of obtaining both the N_f and the N_w results is proportional to the product of the two expressions. In other words the likelihood function in the case of censored tests is given by:

$$L = [1 - \Phi(X_L)]^{N_w} \prod_{i=1}^{m} \varphi(X_i)^{N_i} \qquad (6.13)$$

Similarly for the alternative expression we obtain:

$$L = [1-\Phi(X_L)]^{N_w} \prod_{i=1}^{m} \Delta\Phi_i^{N_i} \qquad (6.14)$$

7. Parameter estimation and choice of interpolating functions

As already mentioned, according to the principle of the maximum likelihood it is presumed that the function which gives the "best fitting" to the experimental results is the one to which corresponds the maximum probability of obtaining these results that is the maximum likelihood.

Suppose that we have obtained a set of results R and that we have selected an interpolating function $\Phi(X;A,B)$. According to the principle of the maximum likelihood, the best estimation of the parameters A,B will be the values A_M,B_M which maximize the likelihood function i.e.:

$$L(R;A_M,B_M) = L_{\max} \qquad (7.1)$$

The same principle can be applied in the choice between different hypotheses about the failure probability function. For example we may have two hypotheses to compare:

Hypothesis 1: $F(X) = \Phi_1(X;A,B)$
Hypothesis 2: $F(X) = \Phi_2(X;C,D)$

Given the set of results R it is possible to calculate the maximum values of the likelihoods of the two hypotheses:

$$L_{\max 1} = L(R;A_M,B_M)$$
$$L_{\max 2} = L(R;C_M,D_M)$$

If $L_{\max 2}>L_{\max 1}$, Φ_2 is to be preferred and vice versa.

8. Test for goodness of fit

As the choice of the interpolating function $\Phi(X;A,B)$ is arbitrary it is necessary to check if the last fits sufficiently well the results. Statistically speaking we have to test the hypothesis H_0, that the failure probability function P(X) or F(X) is equal to the interpolating function $\Phi(X)$ we have chosen, against the hypothesis H_1 that it is not equal to that expression.

For the illustration of this statistical test of hypothesis let us consider first the case of dielectric tests of the type F.

Assuming that H_0 is true, it is possible to calculate the probability P_r of obtaining the results at hand by means of (6.11) which, for facility, is reported here again:

$$P_r = \frac{N!}{N_1!N_2!\ldots N_m!}\prod_{i=1}^{m}\Delta\Phi_i(A,B)^{N_i} \qquad (8.1)$$

The best possible expression for Φ is the one for which (8.1) attains its maximum value. This occurs when the probability $\Delta\Phi_i$, that the result of one test falls in the interval "i", is equal exactly to the observed relative frequency N_i/N i.e. when $\Delta\Phi_i=N_i/N$. Therefore for the best possible expression for $\Phi(x)$, which is not known (and may even not exist), the probability of obtaining the results at hand is :

$$P_{rmax} = \frac{N!}{N_1!N_2!\ldots N_m!}\prod_{i=1}^{m}\left(\frac{N_i}{N}\right)^{N_i} \qquad (8.2)$$

The acceptance or the rejection of the hypothesis H_0 is based on the value of the ratio:

$$g = \frac{P_r}{P_{rmax}} = \frac{\prod_{i=1}^{m}\Delta\Phi_i^{N_i}}{\prod_{i=1}^{m}\left(\frac{N_i}{N}\right)^{N_i}} \qquad (8.3)$$

This ratio is a quantity which clearly lies between 0 and 1. It will be 1 when the hypothesis H_0 is rigorously supported by the results, i.e. when $\Delta\Phi_i=N_i/N$, and tends to 0 as the results N_i/N diverge more and more from the hypothetical values $\Delta\Phi_i$. So, the criterion for rejection or acceptance (in sense of no rejection) of H_0 can be based on a suitably chosen limiting value g_{lim} for the ratio g. In other words we can not reject H_0 if:

$$g \geq g_{lim}$$

The choice of the critical value g_{lim} is made considering the probability distribution of g. In fact, since the results are random g is also a random variable to which corresponds a probability distribution P(g). Wilks [9] has demonstrated that P(g) is independent of the distribution $\Phi(X)$ from which the results are obtained and that -2ln(g) tends asymptotically to the chi-square distribution:

$$-2\ln(g) = \chi^2(\nu) \qquad (8.4)$$

where ν is the number of degrees of freedom which in this case is:

$$\nu = (N_{interv} - 1) - N_{param} \qquad (8.5)$$

N_{interv} being the number of intervals and N_{param} the number of parameters of $\Phi(X)$.

In Appendix 5 it is shown how the cumulative probability distribution of g is obtained from that of χ^2. Having defined the function P(g) we can determine g_{lim} from the condition:

$$P(g_{lim}) = P_{lim} \qquad (8.6)$$

where P_{lim} is the limit we choose for the probability of g being less than g_{lim}. In other words when, from an experiment, we obtain a value of g such that $P(g) \leq P_{lim}$ we do not believe that H_0 is true. Usual values for P_{lim} are 0.1, 0.05 and 0.01 but any other low value can be chosen.

As regards the type P experiments, the statistical test for goodness of fit has a meaning only in the case of "multiple level" (or "up and down") dielectric tests. In fact in such case the results give the failure probability P(U) as function of the stress level and we can make the hypothesis that $P(U)=\Phi(X)$. In order to check that $\Phi(X)$ fits well the experimental results we apply the same procedure described above for the type F tests. It can be summarized in the following steps:

- Calculation of the probability of obtaining the results, assuming that $\Phi(X)$ is true, by expression (6.5) reported here again:

$$P_r = \prod_{i=1}^{k}\frac{N_i!}{N_{fi}!N_{wi}!}P^{N_{fi}}Q^{N_{wi}} \qquad (8.6)$$

- Calculation of the maximum possible value of the probability of obtaining the results, assuming that at each level of the test-voltage U_i the failure probability is equal exactly to the observed frequency of failure i.e. $\Phi(X_j)=N_{fj}/N_j$:

$$P_{rmax} = \prod_{i=1}^{k}\frac{N_i!}{N_{fi}!N_{wi}!}\left(\frac{N_{fi}}{N_i}\right)^{N_{fi}}\left(\frac{N_{wi}}{N_i}\right)^{N_{wi}} \qquad (8.7)$$

- Calculation of the quantities: $g=P_r/P_{rmax}$ and $\chi^2_0=-2\ln(g)$
- Choice of P_{lim} and determination of the corresponding value of g_{lim} as shown in Appendix 6, taking into account that in this case the number of degrees of freedom is:

$$\nu = N_{vlev} - N_{param} \qquad (8.8)$$

where N_{vlev} is the number of test voltage levels.

- Rejection of the hypothesis $P(U)=\Phi(X)$ if $g<g_{lim}$, that is if $P(g)<P_{lim}$, or acceptance if $g \geq g_{lim}$, that is if $P(g)>P_{lim}$.

Since to higher values of P(g) corresponds a better fitting of the curve to the experimental results, the value of P(g) can be considered as an **index of the goodness of fit**.

9. Uncertainties in the estimation of the failure probability function

In the previous sections it was assumed that the failure probability P(U) or F(X) can be represented by the function $\Phi(X;A,B)$. The most likely values of the parameters then were determined maximizing the likelihood function L relative to the type of test under consideration. Finally it was checked if $\Phi(X;A,B)$ fits sufficiently well the experimental results. This analysis, however, is not complete. Since the number of tests is finite, we have to keep in mind that the results obtained are uncertain. If we repeat the experiment we shall obtain, in general, different results and in consequence different values of the parameters. In other words $\Phi(X;A,B)$ is only an estimation of the "true" probability function. It is necessary, therefore, to evaluate the amount of uncertainty of this estimation. For the purpose we have to introduce the concept of confidence regions of the parameters.

9.1 Confidence regions

Let us consider an experiment which is outcome is the set of results R. We make then the hypothesis that the failure probability function is $\Phi(X;A,B)$. Let A_M,B_M be the most likely values of the parameters i.e.:

$$L(R;A_M,B_M) = L_{max} \qquad (9.1)$$

The couple of values A_M,B_M define a point in the parameter space which in this case is the plane {A,B}.

Since the number of results N is finite, the point (A_M,B_M) does not coincide with the point (A_t,B_t) corresponding to the "true values" of the parameters. We must therefore consider the possibility that a certain point, different from (A_M,B_M) may be the "true point" in spite of the fact that its likelihood is lower than L_{max}.

The hypothesis that a point (A,B) may be the "true point" can not be rejected, unless the ratio "a" between the corresponding likelihoods is lower than a limiting value a_{lim}:

$$a = \frac{L(R;A,B)}{L(R;A_M,B_M)} = \frac{L(R;A,B)}{L_{max}} \leq a_{lim} \qquad (9.2)$$

In the vast majority of practical cases, the **likelihood ratio "a"** is particularly well behaved, being a continuous function with a single maximum, away from the extremes of the range of variation of (A,B). In consequence, any line a(A,B) = const. divides the plane {A,B} in two regions, one containing all the points associated with "a" greater, and the other containing those associated with "a" smaller than the chosen constant. It follows that the pair of values A,B which satisfy the equation

$$L(R;A,B) = a_{lim}L_{max} \qquad (9.3)$$

define the boundary between the two regions. This is illustrated schematically in Fig.6.

It is possible to choose the value of a_{lim} in such a way as to obtain a probability C that the point (A_t,B_t), corresponding to the "true values" of the parameters, is contained in the region for which $a > a_{lim}$. In other words if we repeat the same experiment several times we shall obtain, in general, different results each time and therefore different boundaries for that region but, in about 100·C% of the cases, the point (A_t,B_t) will be included in it. The region defined in this way is named **confidence region of the parameters**. The probability C is named **confidence level** and is expressed either in per unit (C) or in per cent (100·C).

The determination of the value of a_{lim} corresponding to the confidence level C is based on a theorem due to S.S. Wilks [10] who demonstrated that for large N the quantity -2ln(a) follows the chi-square distribution that is:

$$-2\ln(a) = \chi^2(\nu) \qquad (9.4)$$

where ν is the number of degrees of freedom which, in this case, is equal to the number of parameters on which $\Phi(X)$ depends:

$$\nu = N_{param} \qquad (9.5)$$

It should be noted that "a" has the same probability distribution as "g", only the number of degrees of freedom is defined in a different way. So, taking into account that $C=1-P_{lim}$, the value of a_{lim} can be determined in the same way as the value of g_{lim} (see Appendix 6).

The value of a_{lim} depends on the confidence level C and the number of parameters N_p. Some values of a_{lim} are given in Table 5. It can be seen that they decrease with the increase of C and N_p.

The area of the confidence region depends on the value of a_{lim} and on the number of tests N. It diminishes when a_{lim} and N increase.

The shape of the "locus", boundary of the confidence region, depends on the type of probability function $\Phi(X)$, on the parameters chosen, on the experimental results R,

TABLE 5   LIMITING VALUES FOR THE LIKELIHOOD RATIO "a"

N_p	C = 50%		C = 70%		C = 90%		C = 95%	
	χ^2_{lim}	a_{lim}	χ^2_{lim}	a_{lim}	χ^2_{lim}	a_{lim}	χ^2_{lim}	a_{lim}
1	0.455	0.797	1.074	0.584	2.706	0.258	3.841	0.147
2	1.386	0.500	2.408	0.300	4.605	0.100	5.991	0.050
3	2.366	0.306	3.665	0.160	6.252	0.044	7.815	0.020
4	3.357	0.187	4.878	0.087	7.779	0.020	9.488	0.009
5	4.351	0.114	6.064	0.048	9.236	0.010	11.070	0.004

etc. In the vast majority of practical cases, the locus is closed as shown in Fig.6. If the locus has a shape similar to a circumference or to an ellipse having axes parallel to the horizontal and vertical ones as in Fig.7a, the parameters A and B are independent. On the contrary, a shape similar to the examples of Fig.7b and 7c is indication of some relation between A and B. In such cases a suitable change of parameters in the expression of Φ(X) may be convenient.

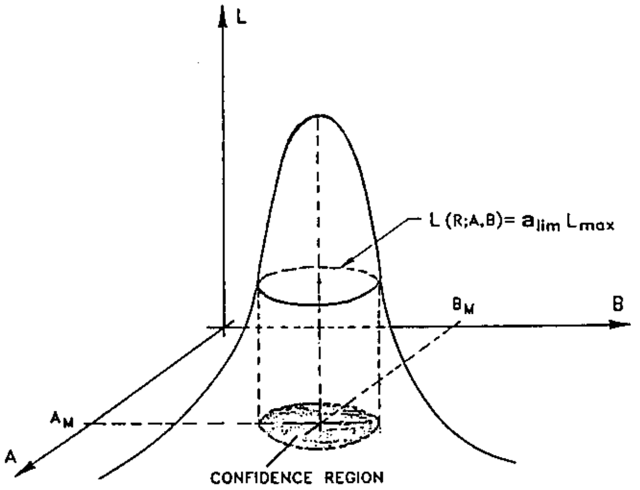


Fig.6 - Representation of the likelihood function and the confidence region of the parameters

When the probability function depends on one parameter only, that is $\Phi=\Phi(X;A)$, the confidence region is reduced to a segment of a straight line as indicated in Fig.8. This segment A_{min}—A_{max} is named also **confidence interval of the parameter A.**

In the case of three parameters i.e. $\Phi=\Phi(X;A,B,C)$ the confidence region becomes a 3-dimensional body like the one shown in Fig.9.

In the general case of N_p parameters the confidence region is a body in the hyper space with N_p dimensions and can not be represented graphically.

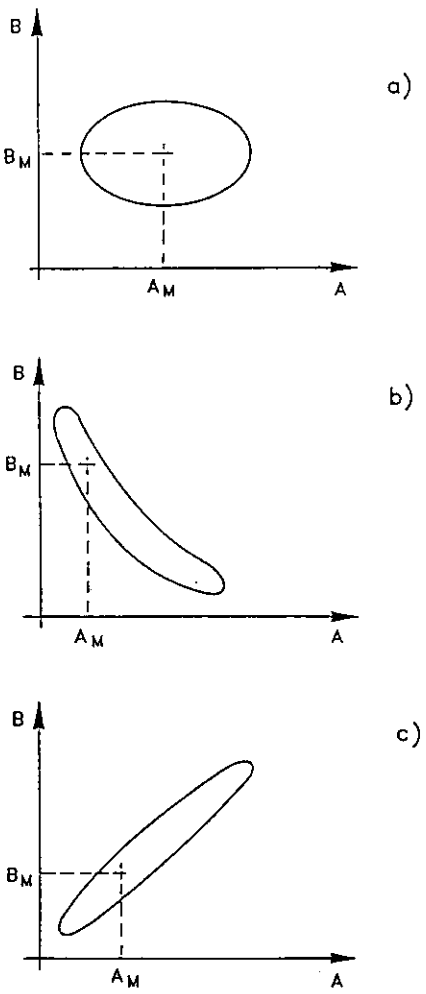


Fig.7 - Examples of confidence regions:
a) Independent parameters
b) and c) Parameters with some relation between them

9.2 Confidence bands for Φ(X)

From the confidence region it is possible to determine a **confidence band** for Φ(X) which contains the true" function Φ with confidence C%. For the purpose it is expedient to determine, for any value of the failure probability Φ, the extreme values

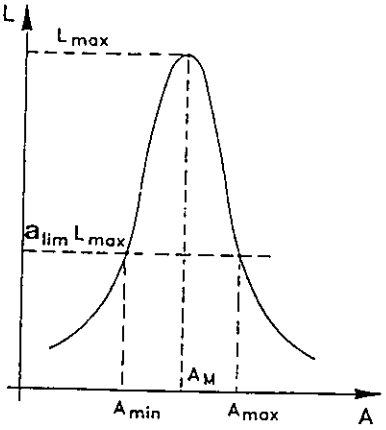


Fig.8 - Determination of the confidence interval of a parameter when the interpolating function depends only on this parameter

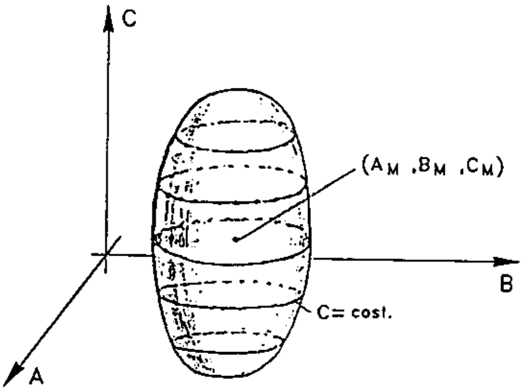


Fig.9 - Example of confidence region when the probability function depends on three parameters

X_{min} and X_{max} that the variable X can take when the values of the parameters vary within the confidence region. In order to obtain X_{min} and X_{max} we have to consider the "limiting values" of the parameters that is the points in the parameter space that lay on the border of the confidence region. Mathematically this means that for $\Phi=\Phi_i$ we have to determine the maximum and the minimum of the function $X(\Phi_i;A,B,...)$ subjecting the values of the parameters A,B,... to the condition:

$$L(\Phi_i;A,B,...) = a_{lim}L_{max} \qquad (9.6)$$

To illustrate better the problem, let us consider again the 2-parameter case. The points located on the boundary of the region represent the "limiting values" of the parameters. For any given value Φ_i of the failure probability, it is possible to calculate the values of the variable $X(\Phi_i;A_j,B_j)$ where the points (A_j,B_j) are approximately uniformly distributed along the boundary of the confidence region as in Fig.10. Between all these calculated values we can select the minimum X_{imin} and the maximum X_{imax}. These operations can be repeated for various values of Φ. The points (X_{imin},Φ_i) and (X_{imax},Φ_i) are points of the left and right confidence limit curves for $\Phi(X;A,B)$ indicated in Fig.11.

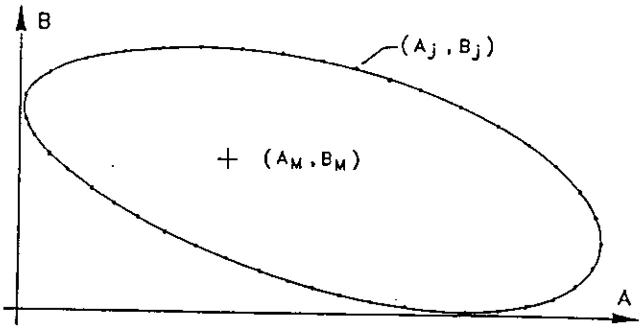


Fig.10 - Points on the boundary of the confidence region used for the determination of the confidence band for the function $\Phi(X)$

For the case of N_p parameters the determination of the confidence band is made following the same procedure but the points in the parameter space, taken for the calculation of X, are nearly uniformly distributed on the surface of the "N_p-dimensional body" representing the confidence region. The number of calculations therefore increases very much with the number of parameters.

The width of the confidence band is diminished increasing the number of tests N or reducing the confidence level C. At constant number of tests it increases with the number of parameters.

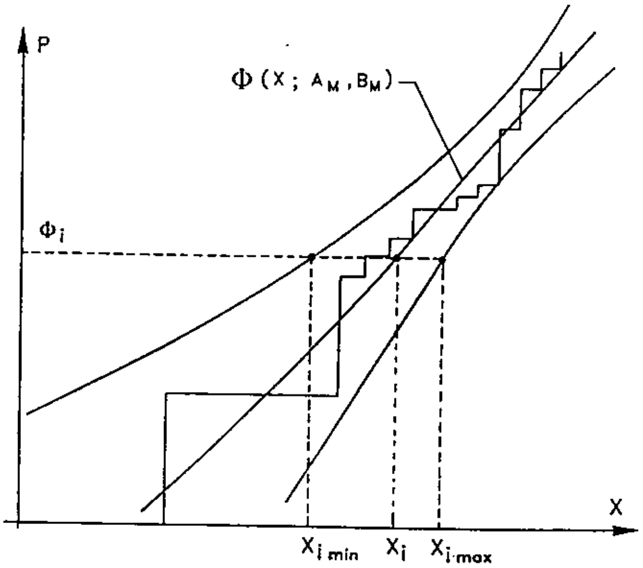


Fig.11 - Confidence band of the interpolating function $\Phi(X)$

9.3 Confidence intervals of the parameters

We have seen that the uncertainty in the estimation of the parameters is represented by their confidence region and by the confidence band for the probability function Φ.

Both representations are graphical. In many cases it is preferred to present synthetically the experimental results in numerical form. The estimated values of the parameters (i.e. their most likely values $A_M, B_M, \ldots$) are frequently used for this purpose. The problem is to represent with a number also the uncertainty of these values.

The first idea that comes to mind is to characterize the uncertainty of each parameter by its **range of variation** determined by the extreme tangents to the boundary R of the confidence region, as shown for two parameters in Fig.12. This characterization is simple but not completely satisfactory. It is preferable to construct for each parameter a confidence interval having a specified, constant probability C of containing the "true" value of the parameter. For the range of variation this probability is higher than the required value C. In fact the tangents define a new, rectangular, confidence region which area is larger than that included in the boundary of the original, "elliptical" one. Obviously to that larger region corresponds a higher confidence level.

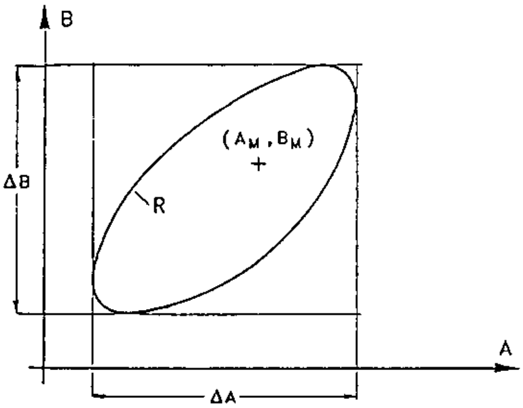


Fig.12 - Range of variation of the parameters

With the purpose to obtain confidence intervals of the required confidence level, it is proposed in [11] to adopt the so called **composite confidence intervals of the parameters**. In order to clarify the meaning of these intervals it is necessary to introduce the concept of **composite hypothesis**.

The procedure we have followed in the determination of the confidence region is based on a statistical testing of hypotheses about the values of the parameters. For example, in the case of 2 parameters we can make the hypothesis H_o: $A=A_h$, $B=B_h$ and test it against the alternative hypothesis H_1: $A \neq A_h$, $B \neq B_h$. The test consists in the calculation of the likelihood ratio:

$$a = \frac{L(R;A_h,B_h)}{L(R;A_M,B_M)} \qquad (9.7)$$

If $a \geq a_{lim}$ we can not reject the hypothesis H_o, if $a < a_{lim}$ we reject it. This kind of hypotheses in which each parameter has a specific value are named **simple hypotheses**.

We can consider a more general case when hypotheses are made about the values of only some of the parameters. For example we can test the hypothesis: H_{oc}: $B=B_h$ without specifying anything about the values of A. Hypotheses of this kind are named **composite hypotheses**. For the testing of these hypotheses we use the **generalized likelihood ratio** which is the ratio between the maximum value of the likelihood of the composite hypothesis H_{oc} and the maximum value of the likelihood of the simple hypothesis about the values of all the parameters. In our example this ratio is:

$$a_c = \frac{\max L(R;A,B_h)}{L(R;A_M,B_M)} \qquad (9.8)$$

N.B. It should be noted that:

$$\max_{A} L(R;A,B_h) = L(R;A^*,B_h)$$

where A^* is the value which maximize $L(R;A,B_h)$ and, in general, it varies with B_h. So, A^* should not be confused with A_M, even if there are cases in which these two values coincide for ex. when $B_h=B_M$ we have $A^*=A_M$.

We can substitute B_h in (9.8) with any other value of B. The hypothesis that one of these values is the "true" value of B can not be rejected unless the corresponding generalized likelihood ratio falls below the critical value a_{clim}. Therefore we can determine an interval for the "possible" values of B from the condition:

$$\frac{L(R;A^*,B)}{L_{\max}} = a_{clim}$$

The values of B obtained from the solution of this equation define the limits of the interval which can be defined as **composite confidence interval for B**.

In the determination of the value of a_{clim} it is necessary to keep in mind that the number of degrees of freedom ν, in the case of composite hypotheses, is equal to the number of parameters with specified values (in our example $\nu=1$).

A handy method for solving (9.9) is to plot in the {A,B} plane the locus defined by the equation:

$$L(R;A,B) = a_{clim} L_{\max} \qquad (9.10)$$

and then to determine the confidence limits B_{min} and B_{max} from the extreme tangents parallel to the A-axis as shown in Fig.13. Obviously the same procedure can be used for the determination of the composite confidence interval for the parameter A. It should be emphasized that the locus defined by (9.10) and indicated with T in Fig.13 is considered here only as a means for the determination of the confidence intervals of the parameters and not as a particular confidence region.

The following considerations can be made about the meaning of the confidence intervals discussed above.

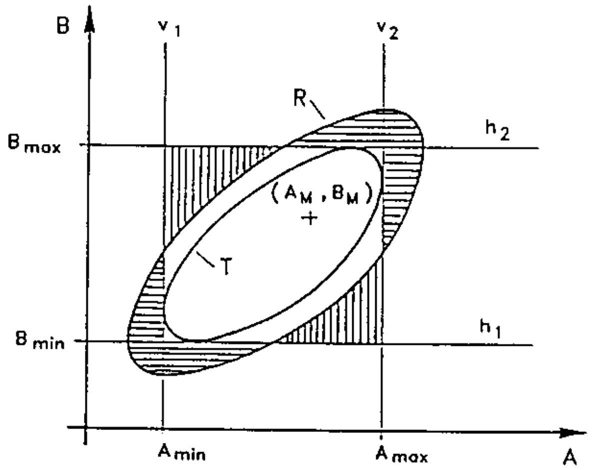


Fig.13 - Determination of the confidence intervals of the parameters

The probability of the event E_A that the abscissa A_t of the "true point" (A_t,B_t) falls in the interval A_{min}—A_{max} is $P(E_A)=C_A$, where C_A is the confidence level we choose. Any point (A,B), contained in the area confined between the two vertical lines v_1 and v_2 in Fig.13 is considered as possible "true point".

Similarly, the probability of the event E_B that the ordinate B_t of the "true point" (A_t,B_t) falls in the interval B_{min}—B_{max} is $P(E_B)=C_B$, where C_B is the confidence level we choose for this interval (usually $C_B=C_A$ but this is not a necessary condition). Any point (A,B), contained in the area confined between the two horizontal lines h_1 and h_2 is considered as a possible "true point".

Since the occurrence of E_A does not depend on whether E_B has occurred or not, the two events are statistically independent. The probability that they occur **simultaneously** is therefore:

$$P(E_A E_B) = C_A C_B = C \qquad (9.11)$$

The rectangular area defined by the intersection of the lines v_1 and v_2 with h_1 and h_2 can be considered as another 100·C% confidence region, to replace the original one defined by equation (9.3), and the intervals A_{min}—A_{max}, B_{min}—B_{max} can be considered as 100·C% confidence intervals relative to the joint event $E_A E_B$.

So, the determination of the composite confidence intervals in the case of two parameters consists in the following steps.

- Calculation of the most likely values of the parameters A_M, B_M in the normal way, maximizing the likelihood function i.e.: $L(A_M,B_M)=L_{max}$.
- Determination of the limits A_{min} and A_{max} of the 100·C_A% composite confidence interval of A by plotting, parallel to the B-axis, the extreme tangents to the locus T_A defined by the expression (9.10). The value of a_{clim} is obtained from Table 5 for $\nu=N_p=1$ and for the selected value of C_A.
- Application of the same procedure to B, that is: determination of the limits B_{min} and B_{max} of the 100·C_B% composite confidence interval of B by plotting, parallel to the A-axis, the extreme tangents to the locus T_B.

N.B. If $C_A=C_B$ the two loci T_A and T_B coincide: $T_A=T_B=T$.

For example, if we want to determine the 95% composite confidence intervals of A and B, we can choose from Table 5 $a_{lim}=0.147$ corresponding to $\nu=N_p=1$ and $C_A=C_B=95\%$.

The confidence intervals determined in this way have double meaning.
When we are interested only in the parameter A, without specifying anything about B, the interval A_{min}—A_{max} is a 100·C_A=95% composite confidence interval. When we are interested simultaneously in the values of both A and B, the same interval A_{min}—A_{max} is to be considered as 100·C_A·C_B=100·0.95²=90% confidence interval.

As regards the rectangular confidence region defined by the intersection of the four tangents, compared with the original one, it has the disadvantage to include points with likelihood lower than that of points which are excluded. This is illustrated in Fig.13 where the areas of the rectangular region, for which $L(R;A,B)<a_{lim}L_{max}$, are dashed with vertical lines and the areas out of the rectangular region, for which $L(R;A,B)>a_{lim}L_{max}$, are dashed with horizontal lines. In spite of this disadvantage, since no better way of specifying confidence intervals for the parameters is available at present, it is suggested to use, when necessary, the confidence intervals obtained from the tangents to the locus T defined by (9.10).

Confidence intervals can be determined not only for the parameters but also for any value of the variable X. For the purpose a confidence band for $\Phi(X)$ is obtained from the locus T (instead of R) by means of the calculations described in Section 9.2

10. Comparison tests

Often the purpose of dielectric testing is to compare two or more insulating structures in order to determine which one is "better". This is necessary, for example:
- in selection of insulating materials
- in selection of a suitable treatment of a type of insulation (such as drying, degassing and impregnation of solid insulation, filtering and degassing of insulating liquids etc.)

- in investigations on the effects of various factors (such as humidity, temperature, ageing etc.) on a given insulation characteristic, etc., etc..

Let us consider the outcome of such "comparison tests" performed on two test objects and denote shortly with R_1 and R_2 the two sets of results obtained. The analysis of these results involves **statistical testing of hypotheses** which consists of the following steps:

- Formulation of two hypotheses: a basic, or working hypothesis, H_0 (named also null hypothesis) and an alternative hypothesis H_1.
- Definition of a test statistic which depends on H_0 and is a function of the experimental results. Its probability distribution must be known.
- Calculation of the value of the test statistic from the experimental results and determination of the probability corresponding to this value.
- Rejection of the hypothesis H_0 and acceptance of H_1 if the above probability is lower than a pre-established limit.

In our case the two hypotheses are:

H_0: The results R_1 and R_2 come from the same probability distribution (i.e. the test objects are equal). The differences between the two sets of results are not significant and can be attributed to statistical fluctuations.

H_1: The results R_1 and R_2 come from different probability distributions (i.e. different test objects). The differences between the two sets of results are significant.

Various test statistics can be used for the testing of these hypotheses. The one suggested by Wilks [9] is defined by the ratio:

$$b = \frac{P^*_{rmax}}{P_{rmax}} \qquad (10.1)$$

where: P^*_{rmax} is the maximum possible value of the probability of obtaining the results R_1 and R_2, subjected to the condition that the set of failure probabilities P_{1i} and P_{2i}, relative to each interval, or test-voltage level, "i" are the same, that is $P_{1i}=P_{2i}=P_i$.

P_{rmax} is the maximum possible value of the probability of obtaining the results R_1 and R_2 without the above restriction.

To make expression (10.1) more clear, let us consider its application for type F tests. For such tests the results R_1 and R_2 can be presented in a table giving the number of failures N_{1i} and N_{2i} which fall in each interval "i" of the X-axis (see Sect. 6.2). With expression (8.2) we can calculate the maximum possible value of the probabilities of obtaining separately the results R_1 and R_2:

$$P_{rmax1} = K_1\prod_{i=1}^{m}\left(\frac{N_{1i}}{N_1}\right)^{N_{1i}}$$
$$P_{rmax2} = K_2\prod_{i=1}^{m}\left(\frac{N_{2i}}{N_2}\right)^{N_{2i}} \qquad (10.2)$$

where: $K_1=N_1!/N_{11}!N_{12}!...N_{1m}!$ and $K_2=N_2!/N_{21}!N_{22}!...N_m!$. The maximum possible value of the probability of obtaining simultaneously the two sets of results R_1 and R_2 is $P_{rmax}=P_{rmax1}P_{rmax2}$ and therefore:

$$P_{rmax} = K_1K_2\prod_{i=1}^{m}\left(\frac{N_{1i}}{N_1}\right)^{N_{1i}}\left(\frac{N_{2i}}{N_2}\right)^{N_{2i}} \qquad (10.3)$$

Let us assume now that the failure probabilities relative to each interval "i" are the same and equal to the relative frequencies observed, taking together all the results R_1 and R_2, that is:

$$P_{1i} = P_{2i} = P_i = \frac{N_{1i} + N_{2i}}{N_1 + N_2} = \frac{N_i}{N} \qquad (10.4)$$

Under this restriction the maximum possible values of the probabilities of obtaining separately the results R_1 and R_2 are respectively:

$$P^*_{rmax1} = K_1\prod_{i=1}^{m}\left(\frac{N_i}{N}\right)^{N_{1i}}$$
$$P^*_{rmax2} = K_2\prod_{i=1}^{m}\left(\frac{N_i}{N}\right)^{N_{2i}} \qquad (10.5)$$

The maximum value of the probability of obtaining simultaneously the two sets of results under the condition (10.4) is $P^*_{rmax}=P^*_{rmax1}P^*_{rmax2}$ and therefore:

$$P^*_{rmax} = K_1K_2\prod_{i=1}^{m}\left(\frac{N_i}{N}\right)^{N_i} \qquad (10.6)$$

In consequence for the ratio b we obtain:

$$b = \frac{\prod_{i=1}^{m}\left(\frac{N_i}{N}\right)^{N_i}}{\prod_{i=1}^{m}\left(\frac{N_{1i}}{N_1}\right)^{N_{1i}}\left(\frac{N_{2i}}{N_2}\right)^{N_{2i}}} \qquad (10.7)$$

This expression can be easily extended to the case of more than two sets of results, say s sets:

$$b = \frac{\prod_{i=1}^{m}\left(\frac{N_i}{N}\right)^{N_i}}{\prod_{j=1}^{s}\prod_{i=1}^{m}\left(\frac{N_{ji}}{N_j}\right)^{N_{ji}}} \qquad (10.8)$$

where:

$$N_i = \sum_{j=1}^{s} N_{ji} \qquad N = \sum_{j=1}^{s} N_j$$

As regards the tests of type P, let us consider s sets of results of tests performed at k voltage levels. For a given voltage level U_i, the results of the set j consist of N_{fji} failures, out of N_{ji} voltage applications, and of $N_{wji}=N_{ji}-N_{fji}$ withstands (see sections 3.1 and 6.1). Taking into account expression (8.7) and proceeding in the same way as indicated above for the type F tests, we obtain :

$$b = \frac{\prod_{i=1}^{k}\left(\frac{N_{fi}}{N_i}\right)^{N_{fi}}\left(\frac{N_{wi}}{N_i}\right)^{N_{wi}}}{\prod_{j=1}^{s}\prod_{i=1}^{k}\left(\frac{N_{fji}}{N_{ji}}\right)^{N_{fji}}\left(\frac{N_{wji}}{N_{ji}}\right)^{N_{wji}}} \qquad (10.9)$$

where:

$$N_{fi} = \sum_{j=1}^{s} N_{fji} \qquad N_{wi} = \sum_{j=1}^{s} N_{wji} \qquad N_i = \sum_{j=1}^{s} N_{ji}$$

Wilks has demonstrated that the ratio "b" has the same probability distribution as "g" and "a" i.e.: $-2\ln(b)=\chi^2(\nu)$. Since no parameters are determined, the number of degrees of freedom in this case is:

$$\begin{aligned} \nu &= m - 1 && \textit{for type F tests} \\ \nu &= k && \textit{for type P tests} \end{aligned} \qquad (10.10)$$

where m is the number of intervals and k is the number of test voltage levels.
It is necessary to emphasize here that the intervals chosen on the X-axis for type F tests, or the test-voltage levels for type P tests, must be the same for both sets of results R_1 and R_2.

Once the value of b is calculated from the results it can be compared with the limiting value b_{lim} as indicated in Appendix 6 for g. If $b \le b_{lim}$ the hypothesis H_0 is to be rejected and vice versa.

It should be noted that in this statistical test no hypotheses were made about the type of distribution of the failure probability. When such hypotheses are made it is possible to use the parameters of the assumed distribution as test statistic. If the confidence regions corresponding to the two sets of results have a **sufficiently large common area** as shown in Fig.14 a), the hypothesis H_0 can not be rejected. In Fig.14 b) is presented a case in which the two confidence regions do not have a common area and H_0 must be rejected in spite of the fact that the confidence intervals of the parameters overlap. This example shows that it is possible to draw wrong conclusions when only the confidence intervals of the parameters are compared individually, without considering the confidence region.

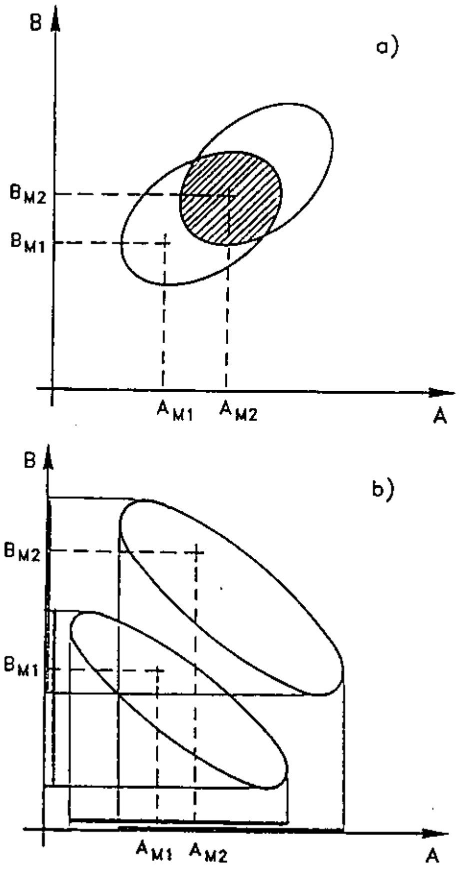


Fig.14 - Confidence regions of two sets of results:

a) The hypothesis H_0 that the two sets of results come from the same population can not be rejected

b) The hypothesis H_0 is to be rejected in spite of the fact that the ranges of variation of the parameters overlap

As mentioned above the hypothesis H_0 can be accepted if there is a **sufficiently large area** common to the two confidence regions. This criterion is rather vague and in many cases it is difficult to decide if the common area is sufficiently large or not. In order to avoid this difficulty, another procedure can be adopted. It consists in the determination of the values of the parameters, taking all the results together, maximizing the likelihood of Φ: $L(R_1+R_2;A,B)$. The goodness of fit of the resulting failure probability curve $\Phi(X;A_M,B_M)$ is then checked for each individual set of results R_1 and R_2. If we found $g<g_{lim}$, for at least one of the two sets of results, the hypothesis H_0 is to be rejected.

- • -

At this point it is interesting to note the uniform approach to the solution of the main problems considered in this paper. This approach is based on the testing of statistical hypotheses. In fact, it was shown that hypotheses about the type of failure probability function, the values of its parameters and, in the case of comparison, the significance of the difference between two or more sets of results, are accepted or rejected according to the value of the ratios, calculated from the experimental results:

g - in the tests for goodness of fit
a - in the determination of the confidence region of the parameters
b - in the comparison tests

All these ratios are random variables and their logarithm tends to the chi-square distribution with number of degrees of freedom specific for each case. Limiting values for g, a and b are then obtained from the corresponding distributions. If the ratios, found from the tests, fall below the limiting values the correspondent hypotheses are rejected.

Since the probability distributions of g, a and b have been demonstrated for large number of results ($N \to \infty$), the validity of their application in the case of small number of results (say $5 \le N \le 20$) is to be checked by further studies.

11. Computer programs

The calculations necessary for the statistical analysis described in sections 5 to 10 require the utilization of digital computers. An example of a computer program suitable for such calculations is briefly illustrated in this section.

The most important features of the program can be seen from the simplified flow-chart shown in Fig.15

- **The main menu** gives four choices:
  - Data file creation or modification of existing data files.
  - Folder management (operations on the archives i.e. copy files from one folder to another, delete files etc.).
  - Calculations.
  - End of job.

- To start the calculations the user must give the following information:
  - The **type of analysis** i.e. if data of a single dielectric test are to be analyzed or data of two or more dielectric tests are to be compared in order to check the hypothesis that all the sets of results come from the same probability distribution (comparison test).
  - The **file containing the data** to analyze. In the case of comparison tests at least two files must be indicated.
  - The **type of test** (multiple level, ramp, stair, constant voltage etc.). With this indication the expression of the likelihood function is automatically selected: expr. (6.6) for tests of type P, expr. (6.9) or (6.11) for type F tests. If the last are censored, expr. (6.13) or (6.14) is used.
  - The **interpolating function** Φ (Gauss, Gumbel, Weibull etc.).
  - The **confidence level** C to be used for the construction of the confidence region and that for the confidence intervals of the parameters.
  - The **choice of the diagrams to be plotted**. There are options for the plotting of several diagrams on the same drawing in order to facilitate comparisons.

- The hart of the program are the two main routines indicated in the flow-chart with "determination of parameters" and "confidence region, intervals and band".
  The first routine finds the most likely values of the parameters by maximization of the likelihood function L.
  The test for goodness of fit is then performed i.e. a statistical test of the hypothesis H_0 that the probability function P(U) or F(X) is equal to the selected interpolating function $\Phi(X)$ (see Section 8). If H_0 is rejected another expression of $\Phi(X)$ must be selected and the calculations are repeated. If H_0 is accepted the calculations continue with the second routine which determines the boundary of the confidence region, the confidence band of $\Phi(X)$ and the confidence intervals of the parameters (see Section 9).
- There are two possibilities for the comparison test: to use the ratio b suggested by Wilks or to check if the curve $\Phi(X)$, relative to all the sets of results taken together, fits well each single set of results (see Section 10).

The program described above has been used for the analysis of the results of examples of the various types of tests discussed in the next section.

12. Examples and comments

In Appendix 7 are given examples of statistical analysis applied to the various types of dielectric tests listed in Table 1. Comments are made here on each type of test and the corresponding example. Some of the comments are specific for the type of test considered while others are general, applicable to all types of dielectric tests.

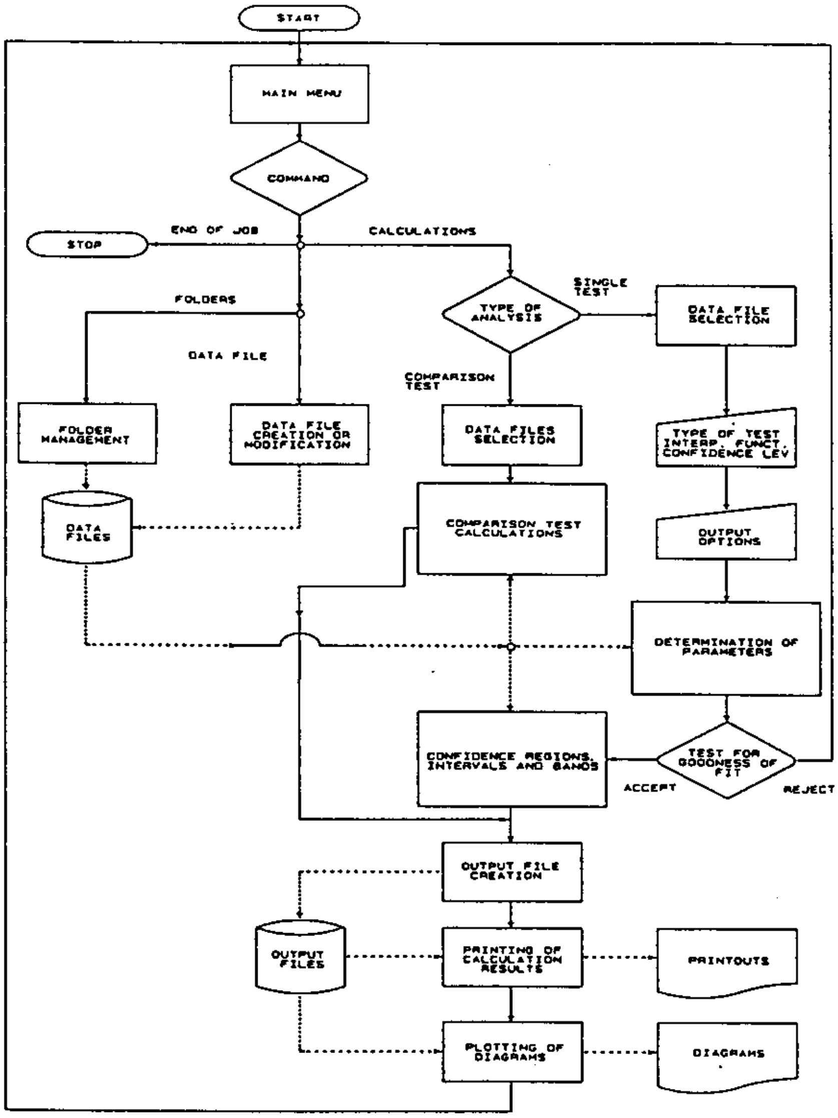


Fig.15 - Simplified flow-chart of the computer program used for the statistical analysis of the results of the various types of dielectric tests given in Appendix 7

Multiple level test - Example Nr 1

The multiple level test is a typical test of type P. In fact, it can be seen from the table of experimental results that increasing the voltage level not always leads to higher number of failures. For instance at 835 kV there are no failures while at 815 kV there are 2 failures. At 950 kV there are 18 failures while at 930 kV there are 19. These irregularities reflect the uncertainty of the estimation of the failure probability from a limited number of test-voltage applications. Such decrease of the observed frequency of failure is impossible in the case of type F tests where the cumulative frequency of the results can only increase with the voltage.

In this example two different expressions for the interpolating function $\Phi(U)$ have been considered:

a) Modified Weibull with $K_0=4$ i.e. $\alpha=4.83$ and therefore:

$$\Phi_a = 1 - 0.5^y \qquad y = \left(1 + \frac{U - U_{50}}{4Z}\right)^{4.83}$$

b) Modified Weibull with $K_0=2$ i.e. $\alpha=2.00$ and therefore:

$$\Phi_b = 1 - 0.5^y \qquad y = \left(1 + \frac{U - U_{50}}{2Z}\right)^{2}$$

The test for goodness of fit shows that, in the case a) we have $P(g)=0.238>0.1$ and in the case b) $P(g)=0.0736<0.1$. Therefore, if we choose $P_{lim}=0.1$, the hypothesis that $P(U)=\Phi_a$ can be accepted while that of $P(U)=\Phi_b$ is to be rejected. In fact, from the diagrams of P(U) in Fig.1(E1) and Fig.2(E1) it can be seen that the number of experimental points located out of the confidence band R are 2 in the case a) and 5 in the case b).

In Fig.1(E1) are shown the two loci R and T as well as the corresponding confidence bands of P(U) for the case a). It is expedient to repeat here the meaning of these two confidence bands.

For the confidence band R is valid the statment that there is 90% probability for the "true curve P(U)" to be contained in it.

The confidence band T is a means for the determination of the 90% confidence intervals of U(P). For example it can be seen from the figure that:

$$U(10\%) = 835 \text{ kV}$$
$$U(10\%)_{min} = 818 \text{ kV}$$
$$U(10\%)_{max} = 848 \text{ kV}$$

Up and down tests - Examples Nr 2, 3 and 4

The most frequently used up and down test is that for the determination of the voltage to which corresponds 50% probability of failure (U_{50}) by the withstand procedure (see Appendix 1). According to this procedure the voltage level is increased by ΔU if no failure occurs after the application of 1 test-voltage or decreased by ΔU in case of failure. Example 2 in Appendix 7 is an illustration of such a test. From Fig.1(E2) it can be seen that the test is suitable for the determination of U_{50} because the minimum width of the confidence band of P(U) is around the probability of 50%.

For the determination of voltage levels with failure probabilities lower than 50% it is necessary to apply not 1 but m test-voltages each time and increase the level by ΔU if no failure occurs or decrease it by ΔU if one or more failures occur in m test-voltage applications. The number m of test-voltages to apply depends on the probability we choose and can be obtained from expression (A1.5) given in Appendix 1.

For instance, if we are interested in the voltage level with 5% probability of failure (i.e.P=0.05) we find m=13.5 and since m must be an integer we have to choose between m=13 to which corresponds P=5.2 % and m=14 to which corresponds P=4.8%.

Once the value of m is defined it is necessary to choose between the "complete" and the "short" procedures. With the "complete" procedure the test-voltage level is decreased by ΔU if one or more failures occur in m applications. With the short procedure the decision to decrease the test-voltage level is taken immediately after the first failure so that the number of applied test-voltages actually is $m_L \leq m$. In this way test time is saved but less information is obtained.

Examples Nr 3 and 4 are an illustration of the short and the complete procedures used for the determination of U_5 (P=5%) on the same test object of Example 2. The results of Example 3 are obtained from the "complete set" given in the table of experimental results of Example 4 but considering, for each test level, only the test-voltage applications until the first failure.

Comparing the total number of impulses applied in the two cases, it can be seen from the tables that it is 113 with the short and 169 with the complete procedure. In other words about 33% of the test time is saved with the short procedure. On the other hand, the data corresponding to the short procedure is not sufficient for the determination of the 90% confidence region (which in this case expands to infinity). However, it is possible to take advantage from previous experience. For example it is known that for impulse breakdown in air the values for Z are usually of the order of 5-6% of U_{50}. So, we can not consider as possible all the values for Z from 0 to ∞. We can select a more "reasonable" range from 0 to the very large limit of say 20% of U_{50}. The confidence region obtained in this way is shown in Fig.1(E3).

In general, considerations about the "possible" values of all the parameters can be used to establish reasonable ranges and reduce in this way the area of the confidence region.

If there is no previous experience we have to accept the more limited information, that can be obtained from the short procedure data, represented for instance by the 50% confidence region. This region is also shown in Fig.1(E3).

The effect of the additional information obtained from the complete procedure is shown in Fig.1(E4) where the diagrams of the two procedures are compared. It can be seen that the 90% confidence region of the complete procedure, obtained without limitations on the values of the parameters, is smaller than that of the short procedure. In consequence also the corresponding confidence band of P(U) is more narrow.

These examples show very well how the uncertainty of the information obtained from the experiments varies with the number of the applied test-voltages and with the previous experience. This is not possible with the classical methods of fitting a function to test data, for instance with the least square method (at least as it is normally used).

Comparing the absolute values of the parameters, estimated from the data of the three experiments on the same test object, it can be observed that the values of U_{50M} of examples 3 and 4 are significantly lower than that of example 2. This is probably due to the fact that the experiments have been performed in different days at different atmospheric conditions. The last have a significant effect on the impulse breakdown in air. As regards the values of Z_M the differences observed between the three cases are not significant and the hypothesis that Z has the same value in all the three experiments can not be rejected.

Ramp test - Example Nr 5

The ramp test is the classic example of type F test. Since the failure probability is a continuous function of the voltage, the two expressions of the likelihood function (6.9) and (6.11) are both applicable. In Example Nr 5 the analyses made with these two functions are compared. It can be seen that the confidence region corresponding to expression (6.9) is smaller and therefore the confidence band of P(U) is narrower. Expression (6.9) therefore seems to be preferable. However, the differences are unimportant for the 20 results examined and other calculations, not reported here, have shown that for greater number of results they are negligible. Preference then may be given to expression (6.11) because it is valid both for continuous and discrete probability functions and therefore remains the same for all the tests of type F including the stair tests.

The example of F test given in Section 3.2 for the illustration of the graphical method is the same as the Example 5 of Appendix 7. The diagram of F(U) in Fig.4 is then comparable to that of Fig.1(E5). In Fig.16 the two diagrams are compared on the same scale. It can be seen that the confidence band determined with the graphical method (using the table of the Kolmogorov statistic) is significantly larger than that obtained with the statistical method (calculations described in Section 9.2). This is due probably to the fact that no hypotheses are made about the type of failure probability distribution when the Kolmogorov statistic is used, while this distribution is assumed to be known in the determination of the confidence band with the statistical method. The uncertainties related to the unknown type of distribution in the case of application of the graphical method are reflected in the larger confidence band.

Constant voltage test

The constant voltage test is type F test and the statistical analysis of the results is exactly the same as that of the ramp test. For this reason no examples of this test are given here.

Stair test - Example Nr 6

The stair test is also type F test and, if the probability of failure is expressed as function of time, the statistical analysis is again exactly the same as that of the ramp test. However, the usual way of presenting the results is to give the voltage levels at which failures have been observed. As already explained in Section 2.3 F), the failure probability is discrete function of the voltage and, therefore, only expression (6.11) for the likelihood function is applicable.

In Fig.2(E6) the discrete functions $F(U_j)$ and their confidence bands are represented conventionally with continuous lines because otherwise the graphical representation is not sufficiently clear.

It is very important to keep in mind that the curve $F(U_j)$ is valid only for the particular stair voltage with which the results have been obtained. The conventional continuous line can not be used for the estimation of the failure probabilities of "intermediate voltage levels" because such levels can be obtained only changing the characteristics of the stair for example the starting voltage level U_1, the voltage step ΔU, or both U_1 and ΔU. This means that the test-voltage is changed and, in general, to different test-voltages correspond different probabilities of failure even if the voltage level considered is the same.

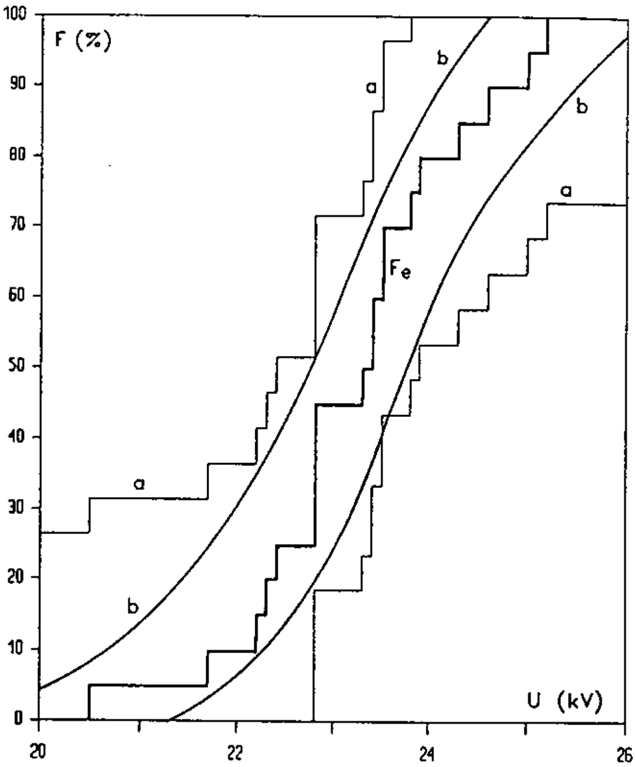


Fig.16 - Comparison of confidence bands obtained with:

a) the graphical method (Sect. 3.2)

b) the maximum likelihood method assuming that F(U) follows a Weibull distribution with $K_0=4$

In Example 6 a set of 100 results is compared with the subset of its first 20 results with the purpose to show the effect of the number of results on the uncertainty of the estimations. From Fig.1(E6) and Fig.2(E6) it can be seen how much the increased number of results reduces the confidence region of the parameters and therefore the confidence band of $F(U_j)$.

Impulse tests performed with the step method Example Nr 7

As indicated in Section 2.3 G), the results of impulse tests performed with the step method can be analyzed in two ways:

A) The test voltage is considered as consisting of one single impulse. The characteristic P(U), i.e. the probability of failure, when one impulse is applied to the test object, is estimated as function of the peak voltage.

B) The test voltage is considered as a series of impulses with increasing peak values (complex stair). The characteristic $F(U_j)$, i.e. the probability of failure within the level U_j, is estimated.

An illustration of these two ways of analysis is given in Example Nr 7. The difference between the two characteristics and between the confidence regions of their parameters can be seen from Fig.1(E7) and Fig.2(E7).

Comparison test - Example Nr 8

In Section 10 it was shown how the testing of statistical hypotheses can be used to answer the question: are the differences between two, or more, sets of results significant or not. Two statistical tests were described there - one based on the ratio b and the other on the ratio g (test for goodness of fit). The application of these two tests is illustrated in detail in Example 8 (Appendix 7). Both tests arrive to the same conclusion that the difference between the two sets of results is significant.

Comparing the two statistical tests it should be noted that the test based on the ratio b requires less calculations and no hypotheses about the failure probability distribution. So, it is to be preferred.

The experimental results in Example 8 have been obtained from ramp tests, which are type F tests, and, in the particular case considered, are not censored. It is possible therefore to calculate the mean value of the failure voltages and their standard mean deviations. Then the classical comparison of two means can be made as indicated in ISO Standard Nr 2854-1976. The outcome of this comparison is that the hypothesis of equality of the two means is to be rejected which is a further confirmation of the conclusion of the two statistical tests illustrated in Example 8.

In Fig.1(E8) are shown the 90% and the 70% confidence regions of the parameters relative to the two sets of results R_1 and R_2. It can be seen that the 90% confidence regions have a common area and it is difficult to decide if it is "sufficiently large" to accept the hypothesis of equality of R_1 and R_2. On the contrary, the 70% confidence regions are completely disjoint and there are no doubts about the rejection of the hypothesis of equality (at the 70% confidence level).

13. Conclusions

Some of the conclusions that can be drawn from this paper are:

1) The response of the insulation to the application of a test-voltage is determined by the characteristics P(U) or F(X) defined in Section 2.
Usually the purpose of a dielectric test is to estimate one of these characteristics and, therefore, from this point of view, the tests are classified in two types: type P and type F.

2) The concept of "test-voltage" was introduced in the paper (see the definition in Section 2.1) because both characteristics refer to a given test-voltage and, in general, they change if the test-voltage is changed.

3) The estimation of the insulation characteristics consists in the determination of a curve which gives "the best fitting" to the experimental results and of its confidence band which reflects the uncertainty due to the finite number of results.

4) The most simple method of estimation is the graphical, which practically does not require calculations, but it has the disadvantage of the subjective selection of the "best fitting curve".

5) The statistical method based on the principle of the maximum likelihood is suggested for general use because of its particular features:

   - The method is general i.e. applicable to all kinds of dielectric tests (including the censored) and to any type of interpolating function.
   - A clear statement of all the hypotheses is required.
   - The uncertainties of the information obtained from the tests is put well in evidence.
   - The "best fitting curve" is defined on the basis of the objective criterion of the maximum likelihood.
   - The hypothesis made about the type of interpolating function permits the determination of confidence bands which are narrower than those obtained with the graphical method (where no hypotheses are made about the type of interpolating function).

6) The disadvantage of the method is that it requires quite complex calculations. However, with the wide diffusion of the digital computers this drawback is to be considered of minor importance.

APPENDIX 1

"UP AND DOWN" TESTS

These tests are used mainly with impulse voltages for tests on self restoring insulation. A description of these tests is given in [5] as follows:

"In an up-and-down test n groups of m substantially equal voltage stresses are applied at voltage levels U_i. The voltage level for each succeeding group of stresses is increased or decreased by a small amount ΔU according to the result of the previous group of stresses.

Two testing procedures are commonly used. **The withstand procedure** aimed at finding voltage levels corresponding to low disruptive discharge probabilities and **the discharge procedure**, which finds voltage levels corresponding to high disruptive discharge probabilities. In the withstand procedure, the voltage level is increased by an amount ΔU if no disruptive discharge occurs in a group of m voltage applications, otherwise the voltage level is decreased by the same amount. In the discharge procedure the voltage level is increased by ΔU if one or more withstands occur, otherwise it is decreased by the same amount. Where m=1, the two procedures become identical and correspond to the well known up-and-down 50% disruptive discharge voltage test.

Tests with other values of m are also used to determine voltages corresponding to other disruptive discharge probabilities. The results are the numbers k_i of stress groups applied at the voltage levels U_i. The first level U_1 taken into account is that at which at least two groups of stresses were applied. The total number of useful groups is $n=\Sigma k_i$".

It is necessary to specify that in the above text with the term "low disruptive discharge probabilities" are indicated probabilities which are lower than 50% and with the term "high disruptive discharge probabilities" the ones which are higher than 50%.

The discharge procedure is applicable for example to tests performed on surge protective devices. In such cases we are interested in the determination of the protection level i.e. in the test voltage level that, when applied to the device, will certainly cause discharge. For this reason, if applying m test voltages of a given level we observe one or more withstands we have to increase the test voltage level. On the contrary, if only discharges occur it is necessary to decrease the voltage level.

More frequently we are interested in the withstand possibilities of the insulation or, more precisely, in the determination of a voltage level to which corresponds a given "low probability" of failure. In such cases the withstand procedure is applicable and when one or more discharges occur in m voltage applications it is necessary to decrease the voltage level. It is to be noted that the decision to decrease the voltage level can be taken immediately after the occurrence of the first discharge without the application of all the m test voltages. With this **short procedure** the number of voltages applied at each level is $m_L \leq m$ and therefore test time can be saved but less information is obtained. In consequence, in many cases, the **complete procedure**, i.e. $m_L = m$ is to be preferred (see the comments on examples Nr 3 and 4 in Section 12).

Further in reference [5] it is indicated that:

"An up and down test provides an estimate of U_p, the voltage at which the disruptive discharge probability is P. U_p^*, the estimate of U_p, is given by:

$$U_p^* = \sum \frac{k_i U_i}{n} \qquad \text{(A1.1)}$$

where k_i is the number of groups of stresses applied at the voltage level U_i.
For a more accurate formula see the technical literature.

To avoid appreciable errors, the lowest voltage level taken into account should not differ from U_p^* by more than $2\Delta U$.

The withstand procedure described above provides an estimate of U_p for a disruptive discharge probability P given by:

$$P = 1 - 0.5^{\frac{1}{m}} \qquad \text{(A1.2)}$$

while the discharge procedure gives U_p for:

$$P = 0.5^{\frac{1}{m}} \qquad \text{(A1.3)}$$

The values of P for which U_p can be estimated in up and down tests are limited by the requirement that m be an integer. Examples are given in the table below".

m	P withstand procedure	P discharge procedure
1	0.50	0.50
2	0.30	0.70
3	0.20	0.80
4	0.15	0.85
7	0.10	0.90
14	0.05	0.95
34	0.02	0.98
70	0.01	0.99

The determination of other parameters of the curve interpolating the results of up and down tests is not recommended in [5].

N.B. The expression (A1.2) is obtained on the basis of the following considerations.

If the probability of failure when one test-voltage is applied at a given voltage level is P, the probability P_m of at least one failure in m test-voltage applications will be:

$$P_m = 1 - (1 - P)^m \qquad \text{(A1.4)}$$

If we choose P_m=50% we obtain the expression (A1.2) and therefore for m we have:

$$m = \frac{\ln(0.5)}{\ln(1 - P)} = - \frac{0.693}{\ln(1 - P)} \qquad \text{(A.5)}$$

The values of m given in the table above are obtained from this last expression.

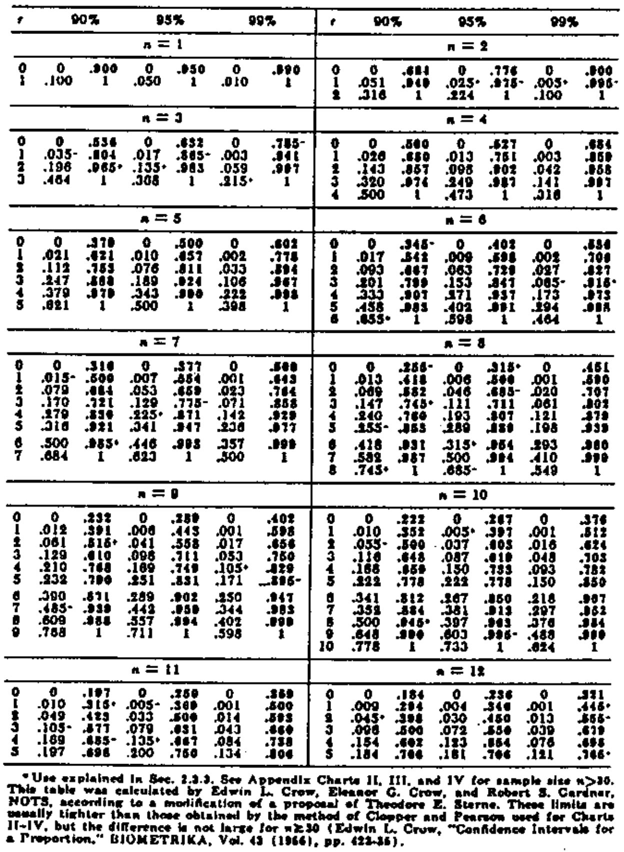
APPENDIX 2

TABLE FOR CONFIDENCE LIMITS OF A PROPORTION

TABLE 21. CONFIDENCE LIMITS FOR A PROPORTION

(a) TWO-SIDED LIMITS*

Upper limits are in boldface. The observed proportion in a random sample is r/n.



* Use explained in Sec. 2.3.3. See Appendix Charts II, III, and IV for sample size n>30. This table was calculated by Edwin L. Crow, Eleanor G. Crow, and Robert S. Gardner, NOTS, according to a modification of a proposal of Theodore E. Sterne. These limits are usually tighter than those obtained by the method of Clopper and Pearson used for Charts II-IV, but the difference is not large for n≥30 (Edwin L. Crow, "Confidence Intervals for a Proportion," BIOMETRIKA, Vol. 43 (1956), pp. 423-35).

TABLE 21 (Contd.)

(a) TWO-SIDED LIMITS (CONTD.)

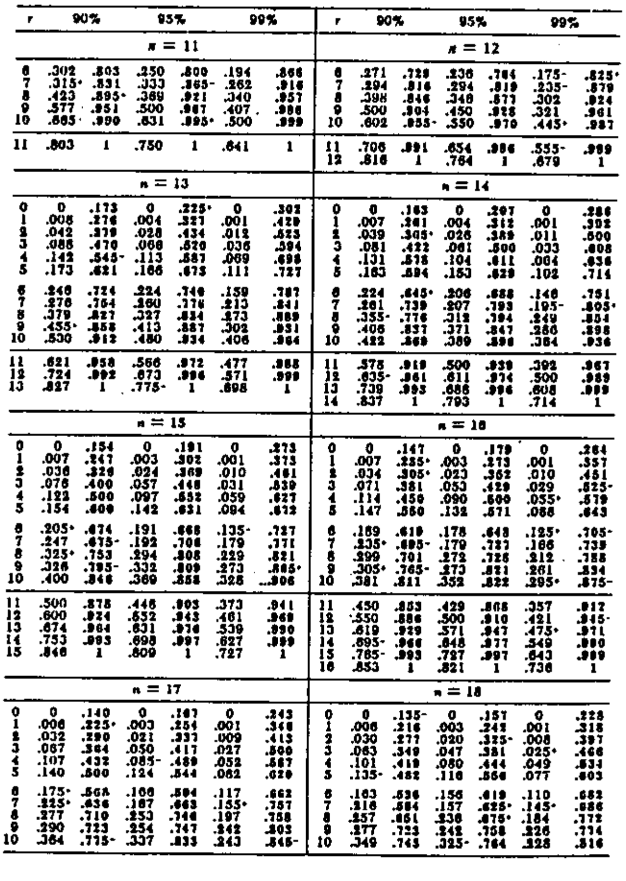


TABLE 21 (Contd.)

(*a*) Two-Sided Limits (Contd.)

r	90%		95%		99%		r	90%		95%		99%	
n = 17							n = 18						
11	.432	.825-	.406	.834	.338	.883	11	.416	.784	.375-	.843	.314	.855+
12	.500	.860	.456	.876	.380	.918	12	.464	.837	.381	.844	.318	.890
13	.568	.893	.511	.915+	.413	.946	13	.518	.865+	.444	.884	.397	.923
14	.636	.933	.583	.950	.500	.973	14	.581	.899	.556	.920	.466	.951
15	.710	.965	.663	.979	.587	.991	15	.651	.937	.619	.953	.534	.975-
16	.775-	.994	.746	.997	.654	.999	16	.723	.970	.675+	.980	.603	.992
17	.860	1	.833	1	.757	1	17	.784	.994	.758	.997	.682	.999
							18	.865+	1	.843	1	.772	1
n = 19							n = 20						
0	0	.130	0	.150	0	.218	0	0	.126	0	.143	0	.209
1	.006	.209	.003	.232	.001	.305+	1	.005+	.203	.003	.222	.001	.293
2	.025	.265+	.019	.316	.008	.383	2	.027	.255-	.018	.294	.008	.375-
3	.059	.337	.044	.365-	.024	.455+	3	.056	.328	.042	.351	.023	.424
4	.095+	.387	.075+	.426	.046	.516+	4	.090	.367	.071	.411	.044	.500
5	.130	.440	.110	.500	.073	.564	5	.126	.422	.104	.467	.069	.576
6	.151	.500	.147	.574	.103	.617	6	.141	.500	.140	.533	.098	.601
7	.209	.613	.150	.635+	.137	.695-	7	.201	.578	.143	.589	.129	.637
8	.238	.614	.222	.655+	.173	.707	8	.221	.633	.209	.649	.163	.707
9	.265+	.663	.232	.688	.212	.782	9	.255+	.642	.222	.706	.200	.726
10	.337	.735-	.312	.768	.218	.788	10	.325-	.675+	.293	.707	.209	.791
11	.366	.762	.345-	.778	.293	.827	11	.358	.745+	.294	.778	.274	.800
12	.387	.791	.365-	.850	.305+	.863	12	.367	.779	.351	.791	.293	.837
13	.440	.849	.426	.853	.383	.897	13	.422	.799	.411	.857	.363	.871
14	.560	.870	.500	.890	.436	.927	14	.500	.859	.467	.860	.399	.902
15	.613	.905-	.574	.925-	.485-	.954	15	.578	.874	.533	.896	.424	.931
16	.663	.941	.635+	.956	.545-	.976	16	.633	.910	.589	.929	.500	.956
17	.735-	.972	.684	.981	.617	.992	17	.672	.944	.649	.958	.576	.977
18	.791	.994	.768	.997	.695-	.999	18	.745+	.973	.706	.982	.625+	.992
19	.870	1	.850	1	.782	1	19	.797	.995-	.778	.997	.707	.999
							20	.874	1	.857	1	.791	1
n = 21							n = 22						
0	0	.123	0	.137	0	.201	0	0	.116	0	.132	0	.194
1	.005+	.192	.002	.213	.000	.283	1	.005-	.182	.002	.205+	.000	.273
2	.026	.245-	.017	.277	.007	.347	2	.024	.236	.016	.264	.007	.334
3	.054	.307	.040	.338	.022	.409	3	.051	.289	.038	.326	.021	.396
4	.086	.353	.068	.398	.041	.460	4	.082	.340	.065-	.389	.039	.454
5	.121	.407	.099	.465+	.065+	.534	5	.115-	.393	.094	.424	.062	.505-
6	.130	.458	.132	.506	.092	.591	6	.116	.444	.126	.500	.088	.550
7	.191	.542	.137	.551	.122	.653	7	.181	.500	.132	.576	.116	.604
8	.192	.593	.197	.602	.155-	.661	8	.182	.556	.187	.582	.147	.666
9	.245-	.647	.213	.662	.189	.717	9	.236	.607	.205+	.617	.179	.682
10	.306	.693	.276	.723	.201	.743	10	.289	.660	.260	.674	.194	.727
11	.307	.694	.277	.724	.257	.799	11	.290	.710	.264	.736	.242	.758
12	.353	.755+	.338	.787	.283	.811	12	.340	.711	.326	.740	.273	.806
13	.407	.808	.398	.803	.339	.845+	13	.393	.764	.383	.795-	.318	.821
14	.458	.809	.449	.863	.347	.878	14	.444	.818	.418	.813	.334	.853
15	.542	.870	.494	.868	.409	.908	15	.500	.819	.424	.868	.390	.884
16	.593	.879	.545-	.901	.466	.935-	16	.556	.884	.500	.874	.450	.912
17	.647	.914	.602	.932	.534	.959	17	.607	.885+	.576	.906	.495+	.938
18	.693	.946	.662	.960	.591	.978	18	.660	.918	.611	.935+	.540	.961
19	.755+	.974	.723	.983	.653	.993	19	.711	.949	.674	.962	.604	.979
20	.808	.995-	.787	.998	.717	1.000	20	.764	.976	.736	.984	.666	.993
21	.877	1	.863	1	.799	1	21	.818	.995+	.795-	.998	.727	1.000
							22	.884	1	.868	1	.806	1

TABLE 21 (Contd.)

(*a*) Two-Sided Limits (Contd.)

r	90%		95%		99%		r	90%		95%		99%	
n = 23							n = 24						
0	0	.111	0	.127	0	.187	0	0	.105+	0	.122	0	.181
1	.005-	.174	.002	.198	.000	.265+	1	.004	.165+	.002	.191	.000	.259
2	.023	.228	.016	.255-	.007	.323	2	.022	.221	.015+	.246	.006	.313
3	.049	.274	.037	.317	.020	.386	3	.047	.264	.035-	.308	.019	.364
4	.078	.328	.062	.361	.038	.429	4	.075-	.317	.059	.347	.036	.416
5	.110	.381	.090	.409	.059	.500	5	.105-	.370	.086	.396	.057	.464
6	.111	.431	.120	.457	.084	.571	6	.105+	.423	.115-	.443	.080	.536
7	.173	.479	.127	.543	.111	.580	7	.165-	.448	.122	.500	.106	.584
8	.174	.522	.178	.591	.140	.616	8	.165+	.552	.169	.557	.133	.636
9	.228	.569	.198	.639	.171	.677	9	.221	.553	.191	.604	.163	.638
10	.273	.619	.247	.640	.187	.702	10	.259	.587	.234	.653	.181	.687
11	.274	.672	.255-	.683	.229	.735-	11	.264	.630	.246	.661	.216	.720
12	.328	.726	.317	.745+	.265+	.771	12	.317	.683	.308	.692	.257	.743
13	.381	.727	.360	.753	.298	.813	13	.370	.736	.339	.754	.280	.784
14	.431	.772	.361	.802	.323	.829	14	.413	.741	.347	.766	.313	.819
15	.478	.826	.409	.822	.364	.860	15	.447	.779	.396	.809	.362	.837
16	.521	.827	.457	.873	.420	.889	16	.448	.835-	.443	.831	.364	.867
17	.569	.889	.543	.880	.429	.916	17	.552	.835+	.500	.878	.416	.894
18	.619	.890	.591	.910	.500	.941	18	.577	.895-	.557	.885+	.464	.920
19	.672	.922	.639	.938	.571	.962	19	.630	.895+	.604	.914	.536	.943
20	.726	.951	.683	.963	.614	.980	20	.683	.925+	.653	.941	.584	.964
21	.772	.977	.745+	.984	.677	.993	21	.736	.953	.692	.965+	.636	.981
22	.826	.995+	.802	.998	.735-	1.000	22	.779	.978	.754	.985-	.687	.994
23	.889	1	.873	1	.813	1	23	.835+	.996	.809	.998	.741	1.000
							24	.895-	1	.878	1	.819	1
n = 25							n = 26						
0	0	.102	0	.118	0	.175+	0	0	.098	0	.114	0	.170
1	.004	.159	.002	.185+	.000	.246	1	.004	.152	.002	.180	.000	.235-
2	.021	.214	.014	.238	.006	.305-	2	.021	.200	.014	.230	.006	.298
3	.045-	.255-	.034	.303	.018	.352	3	.043	.247	.032	.283	.017	.342
4	.072	.307	.057	.336	.034	.403	4	.069	.299	.054	.325+	.033	.393
5	.101	.362	.082	.384	.054	.451	5	.097	.343	.079	.374	.052	.442
6	.102	.390	.110	.431	.077	.500	6	.098	.377	.108	.421	.073	.487
7	.158	.432	.118	.475-	.101	.549	7	.151	.419	.114	.465-	.097	.526
8	.159	.500	.161	.525+	.127	.587	8	.152	.460	.154	.506	.122	.562
9	.214	.568	.185+	.569	.155+	.648	9	.209	.540	.180	.542	.149	.607
10	.246	.610	.222	.616	.175+	.656	10	.233	.581	.212	.579	.170	.658
11	.255-	.611	.236	.664	.205+	.695+	11	.247	.623	.230	.626	.195-	.678
12	.307	.640	.296	.683	.245+	.754	12	.299	.657	.282	.675-	.234	.702
13	.360	.693	.317	.704	.246	.755-	13	.342	.658	.283	.717	.235-	.765+
14	.389	.745+	.336	.762	.305-	.795-	14	.343	.701	.325+	.718	.298	.766
15	.390	.754	.384	.778	.342	.825-	15	.377	.753	.374	.770	.322	.805+
16	.432	.786	.431	.815-	.352	.845-	16	.419	.767	.421	.788	.342	.830
17	.500	.841	.475-	.839	.403	.873	17	.460	.791	.458	.820	.393	.851
18	.568	.842	.525+	.862	.451	.899	18	.540	.848	.494	.846	.438	.878
19	.610	.898	.569	.890	.500	.923	19	.581	.849	.535-	.886	.474	.903
20	.638	.899	.616	.918	.549	.946	20	.623	.902	.579	.894	.513	.927
21	.693	.928	.664	.943	.597	.966	21	.657	.903	.626	.921	.558	.948
22	.745+	.955+	.697	.966	.648	.982	22	.701	.931	.675-	.946	.607	.967
23	.786	.979	.762	.986	.695+	.994	23	.753	.957	.717	.968	.658	.983
24	.841	.996	.815-	.998	.754	1.000	24	.791	.979	.770	.986	.702	.994
25	.898	1	.882	1	.825-	1	25	.848	.996	.820	.998	.765+	1.000
							26	.902	1	.886	1	.830	1

TABLE 21 (Contd.)

(*a*) Two-Sided Limits (Contd.)

r	90%		95%		99%		r	90%		95%		99%	
n = 27							n = 28						
0	0	.093	0	.110	0	.166	0	0	.090	0	.106	0	.162
1	.004	.146	.002	.175-	.000	.225-	1	.004	.140	.002	.170	.000	.218
2	.020	.204	.013	.223	.006	.297	2	.019	.201	.013	.217	.005+	.273
3	.042	.239	.031	.270	.017	.332	3	.040	.232	.030	.259	.016	.323
4	.066	.291	.052	.316	.032	.384	4	.064	.284	.050	.307	.031	.365-
5	.093	.327	.070	.364	.050	.419	5	.089	.312	.073	.357	.048	.408
6	.094	.365+	.101	.415-	.070	.461	6	.090	.355-	.098	.384	.068	.449
7	.145+	.407	.110	.437	.093	.539	7	.139	.396	.106	.424	.089	.500
8	.146	.447	.148	.500	.117	.581	8	.140	.435+	.142	.463	.112	.551
9	.204	.500	.175-	.563	.143	.587	9	.197	.473	.170	.537	.137	.592
10	.221	.553	.202	.570	.166	.617	10	.208	.527	.192	.576	.162	.635+
11	.239	.593	.223	.598	.185-	.668	11	.232	.565-	.217	.616	.175+	.636
12	.291	.635-	.269	.636	.224	.702	12	.284	.604	.258	.619	.214	.677
13	.326	.673	.270	.684	.225-	.716	13	.310	.645+	.259	.645+	.218	.727
14	.327	.674	.316	.730	.284	.775+	14	.312	.688	.307	.693	.272	.728
15	.365+	.709	.364	.731	.298	.776	15	.355-	.690	.355-	.741	.273	.782
16	.407	.761	.402	.777	.332	.815+	16	.396	.716	.381	.742	.323	.786
17	.447	.779	.430	.798	.383	.834	17	.435+	.768	.384	.783	.364	.825-
18	.500	.796	.437	.825+	.413	.857	18	.473	.792	.424	.808	.365-	.838
19	.553	.854	.500	.852	.419	.883	19	.527	.803	.463	.830	.408	.863
20	.593	.855-	.563	.890	.461	.907	20	.565+	.860	.537	.858	.449	.888
21	.635-	.906	.585+	.899	.539	.930	21	.604	.861	.576	.894	.500	.911
22	.673	.907	.636	.924	.581	.950	22	.645+	.910	.616	.902	.551	.932
23	.709	.934	.684	.948	.616	.968	23	.688	.911	.643	.927	.592	.952
24	.761	.958	.730	.969	.668	.983	24	.716	.930	.693	.950	.635+	.969
25	.796	.980	.777	.987	.703	.994	25	.768	.960	.741	.970	.677	.984
26	.854	.996	.825+	.998	.775+	1.000	26	.799	.981	.783	.987	.727	.995-
27	.907	1	.890	1	.834	1	27	.860	.996	.830	.998	.782	1.000
							28	.910	1	.894	1	.838	1
n = 29							n = 30						
0	0	.087	0	.103	0	.160	0	0	.084	0	.100	0	.152
1	.004	.135-	.002	.166	.000	.211	1	.004	.130	.002	.163	.000	.206
2	.018	.190	.012	.211	.005+	.263	2	.018	.183	.012	.206+	.005+	.256
3	.039	.225-	.029	.251	.015+	.316	3	.037	.219	.028	.244	.015-	.310
4	.062	.279	.049	.299	.030	.354	4	.059	.266	.047	.292	.028	.345-
5	.086	.303	.070	.340	.046	.397	5	.083	.295-	.068	.325-	.045-	.388
6	.087	.345-	.094	.374	.065+	.438	6	.084	.336	.091	.364	.063	.430
7	.134	.385+	.103	.413	.086	.477	7	.129	.374	.100	.403	.083	.469
8	.135-	.425-	.136	.451	.108	.523	8	.130	.416	.131	.440	.104	.505+
9	.189	.465	.166	.500	.132	.562	9	.182	.455+	.163	.476	.127	.538
10	.190	.500	.184	.549	.157	.603	10	.183	.492	.175+	.524	.151	.570
11	.225+	.537	.211	.587	.165+	.646	11	.219	.524	.205+	.560	.152	.612
12	.276	.575+	.247	.626	.206	.654	12	.265-	.554	.236	.597	.198	.655+
13	.294	.615-	.251	.660	.211	.684	13	.266	.584	.244	.636	.206	.671
14	.303	.655+	.299	.661	.260	.737	14	.295-	.624	.292	.675+	.249	.692
15	.345-	.697	.339	.701	.263	.740	15	.336	.664	.324	.676	.256	.744
16	.385+	.706	.340	.749	.316	.789	16	.376	.705+	.325-	.708	.308	.751
17	.425-	.724	.374	.753	.346	.794	17	.416	.734	.364	.756	.329	.794
18	.463	.775+	.413	.789	.354	.835-	18	.446	.735+	.403	.764	.345-	.802
19	.500	.810	.451	.816	.397	.843	19	.476	.781	.440	.795-	.388	.848
20	.537	.811	.500	.834	.438	.868	20	.508	.817	.476	.825-	.430	.849
21	.575+	.865+	.549	.864	.477	.892	21	.545-	.818	.524	.837	.462	.873
22	.615-	.866	.587	.897	.523	.914	22	.584	.870	.560	.869	.495-	.896
23	.655+	.913	.626	.906	.562	.935-	23	.624	.871	.597	.900	.531	.917
24	.697	.914	.660	.930	.603	.954	24	.664	.916	.636	.909	.570	.937
25	.721	.938	.701	.951	.646	.970	25	.705+	.917	.675+	.932	.612	.955+
26	.775+	.961	.749	.971	.684	.985-	26	.734	.941	.708	.953	.655+	.972
27	.810	.982	.789	.988	.737	.995-	27	.781	.963	.756	.972	.690	.985+
28	.865+	.996	.834	.998	.789	1.000	28	.817	.982	.795-	.988	.744	.995-
29	.913	1	.897	1	.840	1	29	.870	.996	.837	.998	.794	1.000
							30	.916	1	.900	1	.848	1

TABLE 21 (Contd.)

(*b*) One-Sided Limits*

If the observed proportion is *r*/*n*, enter the table with *n* and *r* for an upper one-sided limit. For a lower one-sided limit, enter the table with *n* and *n* − *r* and subtract the table entry from 1.

r	90%	95%	99%	r	90%	95%	99%	r	90%	95%	99%
n = 2				n = 3				n = 4			
0	.684	.776	.900	0	.536	.632	.785+	0	.438	.527	.684
1	.949	.975-	.995-	1	.804	.865-	.941	1	.680	.751	.859
				2	.965+	.983	.997	2	.857	.902	.958
								3	.974	.987	.997
n = 5				n = 6				n = 7			
0	.369	.451	.602	0	.319	.393	.536	0	.280	.348	.482
1	.584	.657	.778	1	.510	.582	.706	1	.453	.521	.643
2	.753	.811	.894	2	.667	.729	.827	2	.596	.659	.764
3	.888	.924	.967	3	.799	.847	.915+	3	.721	.775-	.858
4	.979	.990	.998	4	.907	.937	.973	4	.830	.871	.929
				5	.983	.991	.998	5	.921	.947	.977
								6	.985+	.993	.999
n = 8				n = 9				n = 10			
0	.250	.312	.438	0	.226	.283	.401	0	.206	.259	.369
1	.406	.471	.590	1	.368	.429	.544	1	.337	.394	.504
2	.538	.600	.707	2	.490	.550	.656	2	.450	.507	.612
3	.655+	.711	.802	3	.599	.655+	.750	3	.552	.607	.703
4	.760	.807	.879	4	.699	.749	.829	4	.646	.696	.782
5	.853	.889	.939	5	.790	.831	.895-	5	.733	.778	.850
6	.931	.954	.980	6	.871	.902	.947	6	.812	.850	.907
7	.987	.994	.999	7	.939	.959	.983	7	.884	.913	.952
				8	.988	.994	.999	8	.945+	.963	.984
								9	.990	.995-	.999
n = 11				n = 12				n = 13			
0	.189	.238	.342	0	.175-	.221	.319	0	.162	.206	.298
1	.310	.364	.470	1	.287	.339	.440	1	.268	.316	.413
2	.415+	.470	.572	2	.386	.438	.537	2	.360	.410	.506
3	.511	.564	.660	3	.475+	.527	.622	3	.444	.495-	.588
4	.599	.650	.738	4	.559	.609	.698	4	.523	.573	.661
5	.682	.729	.806	5	.638	.685-	.765+	5	.598	.645+	.727
6	.759	.800	.866	6	.712	.755-	.825+	6	.669	.713	.787
7	.831	.865-	.916	7	.781	.819	.879	7	.736	.776	.841
8	.895+	.921	.957	8	.846	.877	.924	8	.799	.834	.889
9	.951	.967	.986	9	.904	.928	.961	9	.858	.887	.931
10	.990	.995+	.999	10	.955-	.970	.987	10	.912	.934	.964
				11	.991	.996	.999	11	.958	.972	.988
								12	.992	.996	.999
n = 14				n = 15				n = 16			
0	.152	.193	.280	0	.142	.181	.264	0	.134	.171	.250
1	.251	.297	.389	1	.236	.279	.368	1	.222	.264	.349
2	.337	.385+	.478	2	.317	.363	.453	2	.300	.344	.430
3	.417	.466	.557	3	.393	.440	.529	3	.371	.417	.503
4	.492	.540	.627	4	.464	.511	.597	4	.439	.484	.569
5	.563	.610	.692	5	.532	.577	.660	5	.504	.548	.630

---

* Use explained in Sec. 2.3.3. See Sec. 2.3.3 for sample size $n > 30$. This table was computed by Robert S. Gardner, NOTS.

TABLE 21 (Contd.)

(*b*) One-Sided Limits (Contd.)

r	90%	95%	99%	r	90%	95%	99%	r	90%	95%	99%
	$n = 14$				$n = 15$				$n = 16$		
6	.631	.675-	.751	6	.596	.640	.718	6	.565+	.609	.687
7	.695+	.736	.805+	7	.658	.700	.771	7	.625-	.667	.739
8	.757	.794	.854	8	.718	.756	.821	8	.682	.721	.788
9	.815-	.847	.898	9	.774	.809	.863+	9	.737	.773	.834
10	.869	.896	.936	10	.828	.858	.906	10	.790	.822	.875-
11	.919	.939	.967	11	.878	.903	.941	11	.839	.868	.912+
12	.961	.974	.989	12	.924	.943	.969	12	.886	.910	.945-
13	.993	.996	.999	13	.964	.976	.990	13	.929	.947	.971
				14	.993	.997	.999	14	.966	.977	.990
								15	.993	.997	.999
	$n = 17$				$n = 18$				$n = 19$		
0	.127	.162	.237	0	.120	.153	.226	0	.114	.146	.215+
1	.210	.250	.332	1	.199	.238	.316	1	.190	.226	.302
2	.284	.326	.410	2	.269	.310	.391	2	.257	.296	.374
3	.352	.396	.480	3	.334	.377	.458	3	.319	.359	.439
4	.416	.461	.543	4	.396	.439	.520	4	.378	.419	.498
5	.478	.522	.603	5	.455+	.498	.577	5	.434	.476	.554
6	.537	.580	.658	6	.512	.554	.631	6	.489	.530	.606
7	.594	.636	.709	7	.567	.608	.681	7	.541	.582	.655+
8	.650	.689	.758	8	.620	.659	.729	8	.592	.632	.702
9	.703	.740	.803	9	.671	.709	.774	9	.642	.680	.746
10	.754	.788	.845-	10	.721	.756	.816	10	.690	.726	.788
11	.803	.834	.883	11	.769	.801	.855-	11	.737	.770	.827
12	.849	.876	.918	12	.815-	.844	.890	12	.782	.812	.863
13	.893	.915+	.948	13	.858	.884	.923	13	.825-	.853	.897
14	.933	.950	.973	14	.899	.920	.951	14	.866	.890	.927
15	.968	.979	.991	15	.937	.953	.975-	15	.905-	.925-	.954
16	.994	.997	.999	16	.970	.980	.992	16	.941	.956	.976
				17	.994	.997	.999	17	.972	.981	.992
								18	.994	.997	.999
	$n = 20$				$n = 21$				$n = 22$		
0	.109	.139	.206	0	.104	.133	.197	0	.099	.127	.189
1	.181	.216	.289	1	.173	.207	.277	1	.166	.198	.266
2	.245-	.283	.358	2	.234	.271	.344	2	.224	.259	.330
3	.304	.344	.421	3	.291	.329	.404	3	.279	.316	.389
4	.361	.401	.478	4	.345+	.384	.460	4	.331	.369	.443
5	.415-	.456	.532	5	.397	.437	.512	5	.381	.420	.493
6	.467	.508	.583	6	.448	.487	.561	6	.430	.468	.541
7	.518	.558	.631	7	.497	.536	.608	7	.477	.515+	.587
8	.567	.606	.677	8	.544	.583	.653	8	.523	.561	.630
9	.615+	.653	.720	9	.590	.628	.695+	9	.568	.605-	.672
10	.662	.698	.761	10	.636	.672	.736	10	.611	.647	.712
11	.707	.741	.800	11	.679	.714	.774	11	.654	.689	.750
12	.751	.783	.837	12	.722	.755+	.811	12	.695+	.729	.786
13	.793	.823	.871	13	.764	.794	.845+	13	.736	.767	.821
14	.834	.860	.902	14	.804	.832	.878	14	.775+	.804	.853
15	.873	.896	.931	15	.842	.868	.908	15	.813	.840	.884
16	.910	.929	.956	16	.879	.901	.935-	16	.850	.874	.912
17	.944	.958	.977	17	.914	.932	.959	17	.885+	.906	.938
18	.973	.982	.992	18	.946	.960	.978	18	.918	.935+	.961
19	.995-	.997	.999	19	.974	.983	.993	19	.949	.962	.979
				20	.995-	.998	1.000	20	.976	.984	.993
								21	.995+	.998	1.000

TABLE 21 (Contd.)

(*b*) One-Sided Limits (Contd.)

r	90%	95%	99%	r	90%	95%	99%	r	90%	95%	99%
	$n = 23$				$n = 24$				$n = 25$		
0	.095+	.122	.181	0	.091	.117	.175-	0	.088	.113	.168
1	.159	.190	.256	1	.153	.183	.246	1	.147	.176	.237
2	.215+	.249	.318	2	.207	.240	.307	2	.199	.231	.296
3	.268	.304	.374	3	.258	.292	.361	3	.248	.282	.349
4	.318	.355-	.427	4	.306	.342	.412	4	.295-	.330	.398
5	.366	.404	.476	5	.352	.389	.460	5	.340	.375+	.444
6	.413	.451	.522	6	.398	.435-	.505+	6	.383	.420	.488
7	.459	.496	.567	7	.442	.479	.548	7	.426	.462	.531
8	.503	.540	.609	8	.484	.521	.590	8	.467	.504	.571
9	.546	.583	.650	9	.526	.563	.630	9	.508	.544	.610
10	.589	.625-	.689	10	.567	.603	.668	10	.548	.583	.648
11	.630	.665-	.727	11	.608	.642	.705-	11	.587	.621	.684
12	.670	.704	.763	12	.647	.681	.740	12	.625-	.659	.719
13	.710	.742	.797	13	.685+	.718	.774	13	.662	.695-	.752
14	.748	.778	.829	14	.723	.754	.806	14	.699	.730	.784
15	.786	.814	.860	15	.759	.788	.837	15	.735-	.764	.815+
16	.822	.848	.889	16	.795+	.822	.867	16	.770	.798	.845-
17	.857	.880	.916	17	.830	.854	.894	17	.804	.830	.873
18	.890	.910	.941	18	.863	.885+	.920	18	.837	.861	.899
19	.922	.938	.962	19	.895+	.914	.943	19	.869	.890	.923
20	.951	.963	.980	20	.925+	.941	.964	20	.899	.918	.946
21	.977	.984	.993	21	.953	.965+	.981	21	.928	.943	.966
22	.995+	.998	1.000	22	.978	.985-	.994	22	.955+	.966	.982
				23	.996	.998	1.000	23	.979	.986	.994
								24	.996	.998	1.000
	$n = 26$				$n = 27$				$n = 28$		
0	.085-	.109	.162	0	.082	.105+	.157	0	.079	.101	.152
1	.142	.170	.229	1	.137	.164	.222	1	.132	.159	.215-
2	.192	.223	.286	2	.185+	.215+	.277	2	.179	.208	.268
3	.239	.272	.337	3	.231	.263	.326	3	.223	.254	.316
4	.284	.318	.385-	4	.275-	.308	.373	4	.265+	.298	.361
5	.328	.363	.430	5	.317	.351	.417	5	.306	.339	.404
6	.370	.405+	.473	6	.358	.392	.458	6	.346	.380	.445-
7	.411	.447	.514	7	.397	.432	.498	7	.385-	.419	.484
8	.451	.487	.554	8	.436	.471	.537	8	.422	.457	.521
9	.491	.526	.592	9	.475-	.509	.574	9	.459	.494	.558
10	.529	.564	.628	10	.512	.547	.610	10	.496	.530	.593
11	.567	.602	.664	11	.549	.583	.645+	11	.532	.565+	.627
12	.604	.638	.698	12	.585-	.618	.679	12	.567	.600	.660
13	.641	.673	.731	13	.620	.653	.711	13	.601	.634	.692
14	.676	.708	.763	14	.655+	.687	.743	14	.635+	.667	.723
15	.711	.742	.794	15	.689	.720	.773	15	.669	.699	753
16	.746	.774	.823	16	.723	.752	.802	16	.701	.731	.782
17	.779	.806	.851	17	.756	.783	.831	17	.733	.762	.810
18	.812	.837	.878	18	.788	.814	.857	18	.765-	.792	.837
19	.843	.866	.903	19	.819	.843	.883	19	.796	.821	.863
20	.874	.894	.927	20	.849	.871	.907	20	.826	.849	.888
21	.903	.921	.948	21	.879	.899	.930	21	.855+	.878	.911
22	.931	.946	.967	22	.907	.924	.950	22	.883	.902	.932
23	.957	.968	.983	23	.934	.948	.968	23	.911	.927	.952
24	.979	.986	.994	24	.958	.969	.983	24	.936	.950	.969
25	.996	.998	1.000	25	.980	.987	.994	25	.960	.970	.984
				26	.996	.998	1.000	26	.981	.987	.995-
								27	.996	.998	1.000

TABLE 21 (Contd.)

(*b*) One-Sided Limits (Contd.)

r	90%	95%	99%	r	90%	95%	99%	r	90%	95%	99%
	$n = 29$				$n = 30$						
0	.076	.098	.147	0	.074	.095+	.142				
1	.128	.153	.208	1	.124	.149	.202				
2	.173	.202	.260	2	.168	.195+	.252				
3	.216	.246	.307	3	.209	.239	.298				
4	.257	.288	.350	4	.249	.280	.340				
5	.297	.329	.392	5	.287	.319	.381				
6	.335-	.368	.432	6	.325-	.357	.420				
7	.372	.406	.470	7	.361	.394	.457				
8	.409	.443	.507	8	.397	.430	.493				
9	.445+	.479	.542	9	.432	.465+	.527				
10	.481	.514	.577	10	.466	.499	.561				
11	.515+	.549	.610	11	.500	.533	.594				
12	.550	.583	.643	12	.533	.566	.626				
13	.583	.616	.674	13	.566	.598	.657				
14	.616	.648	.705-	14	.599	.630	.687				
15	.649	.680	.734	15	.630	.661	.716				
16	.681	.711	.763	16	.662	.692	.744				
17	.712	.741	.791	17	.692	.721	.772				
18	.743	.771	.818	18	.723	.750	.799				
19	.774	.800	.843	19	.752	.779	.824				
20	.803	.828	.868	20	.782	.807	.849				
21	.832	.855+	.892	21	.810	.834	.873				
22	.860	.881	.914	22	.838	.860	.896				
23	.888	.906	.935-	23	.865+	.885+	.917				
24	.914	.930	.954	24	.891	.909	.937				
25	.938	.951	.970	25	.917	.932	.955+				
26	.961	.971	.985-	26	.941	.953	.972				
27	.982	.988	.995-	27	.963	.972	.985+				
28	.996	.998	1.000	28	.982	.988	.995-				
				29	.996	.998	1.000				

APPENDIX 3

TABLE FOR THE QUANTILES OF THE KOLMOGOROV TEST STATISTIC

Table A14 Quantiles of the Kolmogorov Test Statistic[a]

One-Sided Test	$p = .90$	.95	.975	.99	.995		$p = .90$	.95	.975	.99	.995
Two-Sided Test	$p = .80$	.90	.95	.98	.99		$p = .80$	.90	.95	.98	.99
$n = 1$	.900	.950	.975	.990	.995	$n = 21$	.226	.259	.287	.321	.344
2	.684	.776	.842	.900	.929	22	.221	.253	.281	.314	.337
3	.565	.636	.708	.785	.829	23	.216	.247	.275	.307	.330
4	.493	.565	.624	.689	.734	24	.212	.242	.269	.301	.323
5	.447	.509	.563	.627	.669	25	.208	.238	.264	.295	.317
6	.410	.468	.519	.577	.617	26	.204	.233	.259	.290	.311
7	.381	.436	.483	.538	.576	27	.200	.229	.254	.284	.305
8	.358	.410	.454	.507	.542	28	.197	.225	.250	.279	.300
9	.339	.387	.430	.480	.513	29	.193	.221	.246	.275	.295
10	.323	.369	.409	.457	.489	30	.190	.218	.242	.270	.290
11	.308	.352	.391	.437	.468	31	.187	.214	.238	.266	.285
12	.296	.338	.375	.419	.449	32	.184	.211	.234	.262	.281
13	.285	.325	.361	.404	.432	33	.182	.208	.231	.258	.277
14	.275	.314	.349	.390	.418	34	.179	.205	.227	.254	.273
15	.266	.304	.338	.377	.404	35	.177	.202	.224	.251	.269
16	.258	.295	.327	.366	.392	36	.174	.199	.221	.247	.265
17	.250	.286	.318	.355	.381	37	.172	.196	.218	.244	.262
18	.244	.279	.309	.346	.371	38	.170	.194	.215	.241	.258
19	.237	.271	.301	.337	.361	39	.168	.191	.213	.238	.255
20	.232	.265	.294	.329	.352	40	.165	.189	.210	.235	.252
						Approximation for $n > 40$	$\frac{1.07}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

SOURCE. Adapted from Table of Miller (1956).

[a] The entries in this table are selected quantiles w_p of the Kolmogorov test statistics T, T^+, and T^- as defined by Equation 6.1.1 for two-sided tests and by Equations 6.1.2 and 6.1.3 for one-sided tests. Reject H_0 at the level α if T exceeds the $1-\alpha$ quantile given in this table. These quantiles are exact for $n \leq 40$ in the two-tailed test. The other quantiles are approximations that are equal to the exact quantiles in most cases. A better approximation for $n > 40$ results if $(n + \sqrt{n/10})^{1/2}$ is used instead of $\sqrt{n}$ in the denominator.

APPENDIX 4

MODIFIED EXPRESSIONS FOR THE GUMBEL AND WEIBULL PROBABILITY DISTRIBUTIONS

The classic formulae for the Gumbel and Weibull distributions can be modified in order to express the probability as function of the median $X_{md}=X_{50}$ and the parameter Z. The reasons for this change of parameters and the definition of Z are given in Section 4.

In order to simplify some of the expressions it is convenient to use the withstand probability Q instead of the failure probability Φ. Obviously:

$$Q = 1 - \Phi \qquad (A4.1)$$

The determination of the modified expressions is given below separately for the two distributions.

Gumbel distribution

From the classic expression of the Gumbel distribution given in Table 4 it follows that:

$$Q = e^{-e^{y}} \qquad (A4.2)$$

and therefore:

$$y = \ln[\ln(\frac{1}{Q})] \qquad (A4.3)$$

On the other hand $y=\alpha(X-X_{mo})$ and, taking into account that $Q(X_{50})=0.5$ and $Q(X_z)=0.8413$, the following two equations are obtained from (A4.3):

$$y_{50} = \alpha(X_{50} - X_{mo}) = \ln[\ln(\frac{1}{0.5})] = -0.3665$$

$$y_z = \alpha(X_{50} - X_z - X_{mo}) = $$
$$= \ln[\ln(\frac{1}{0.8413})] = -1.7556$$

From these equations it is possible to determine y and the old parameters α and X_{mo} as functions of the new parameters X_{50} and Z:

$$\alpha = \frac{1.39}{Z}$$

$$X_{mo} = X_{50} + 0.2638\, Z$$

$$y = 0.3667 + 1.39\frac{(X - X_{50})}{Z}$$

The substitution of this last expression of y in (A4.2) yields:

$$Q = 0.5^{\exp(1.39\frac{X - X_{50}}{Z})} \qquad (A4.4)$$

We can consider: $y^{*}=1.39(X-X_{50})/Z$ as a new standardized variable and therefore the final expression of Φ becomes:

$$\Phi(y^{*}) = 1 - 0.5^{\exp(y^{*})} \qquad (A4.5)$$

N.B. For simplicity, in Table 4, the standardized variable is indicated with y instead of y*.

Weibull distribution

From the classic expression for the Weibull distribution, given in Table 4, it can be seen that:

$$Q = e^{-y} \qquad (A4.6)$$

$$y = (\frac{X - X_o}{X_s})^{\alpha} \qquad (A4.7)$$

As indicated in Section 4, the threshold parameter can be presented in the form:

$$X_0 = X_{50} - K_0 Z \qquad (A4.8)$$

The scale parameter X_s can also be presented in function of Z:

$$X_s = K_s Z \qquad (A4.9)$$

where K_s is a constant depending on K_0 and α. In fact if we put $X=X_{50}$ in (A4.7) we obtain $y=(K_0 Z/K_s)^{\alpha}$. On the other hand from (A4.6) it follows that $y=\ln(1/Q)$ and taking into account that $Q(X_{50})=0.5$, we find:

$$K_s = \frac{K_o}{[-\ln(0.5)]^{\frac{1}{\alpha}}} \qquad (A4.10)$$

The substitution of this value of K_s in (A4.9) and of the resulting expression for X_s in (A4.7) yields:

$$y = -\ln(0.5)\left(1 + \frac{X - X_{50}}{K_o Z}\right)^{\alpha}$$

and in consequence:

$$Q = e^{-y} = 0.5^{\left(1 + \frac{X - X_{50}}{K_o Z}\right)^{\alpha}} \qquad \text{(A4.11)}$$

Since to $X_z = X_{50} - Z$ corresponds $Q(X_z) = 0.8413$, from (A4.11) is obtained:

$$\alpha = \frac{1.39}{\ln\left(\frac{K_o}{K_o - 1}\right)} \qquad \text{(A4.12)}$$

Taking the quantity:

$$y^{*} = \left(1 + \frac{X - X_{50}}{K_o Z}\right)^{\alpha} \qquad \text{(A4.13)}$$

as new standardized variable, the final expression for Φ takes the form:

$$\Phi = 1 - 0.5^{y^{*}} \qquad \text{(A4.14)}$$

N.B. For simplicity, also in this case the standardized variable is indicated in Table 4 with y instead of y*.

APPENDIX 5

CONSTRUCTION OF PROBABILITY PAPERS

From Table 4 it can be seen that for the Normal and the Gumbel distributions the standardized variable y is a linear function of X. In such case the probability paper is constructed as follows.

On one axis, usually the horizontal, is plotted the variable X on a linear scale. On the other axis, usually the vertical, is plotted the standardized variate y also on a linear scale. In addition the probability Φ(y) is plotted on a scale parallel to that of y in manner to have for each value of y the corresponding value of Φ(y). In this way straight lines are obtained for Φ(y) plotted against X.

For example let us consider the Gumbel distribution. In Table 4 is given the expression of Φ as function of y. For the construction of the scales it is more practical to express y as function of Φ i.e.:

$$y = \ln[\ln(\frac{1}{1 - \Phi})] \qquad \text{(A5.1)}$$

Then for any given value of the probability Φ we can calculate the value of y, plot it on the scale and write the corresponding value of Φ.

For the Normal distribution it is not possible to express explicitly y as function of Φ. However, in tabulated form, this function is available in the literature. Here, in Table 1 (A5), are given the values of y corresponding to selected values of the probability Φ. For the Normal distribution the values of y are taken from Table A1 of ref.[5]. For the Gumbel distribution they are calculated using expression (A5.1). The resulting scales are shown in Fig.1(A5).

TABLE 1 (A5)

Φ	y or ln(y)	
	NORMAL (GAUSS)	GUMBEL AND WEIBULL
0.999	3.09	1.93
0.99	2.33	1.53
0.95	1.64	1.10
0.90	1.28	0.83
0.80	0.84	0.48
0.70	0.52	0.19
0.632	0.337	0.00
0.60	0.25	-0.09
0.50	0.00	-0.37
0.40	-0.25	-0.67
0.30	-0.52	-1.03
0.20	-0.84	-1.50
0.10	-1.28	-2.25
0.05	-1.64	-2.97
0.01	-2.33	-4.60
0.001	-3.09	-6.91

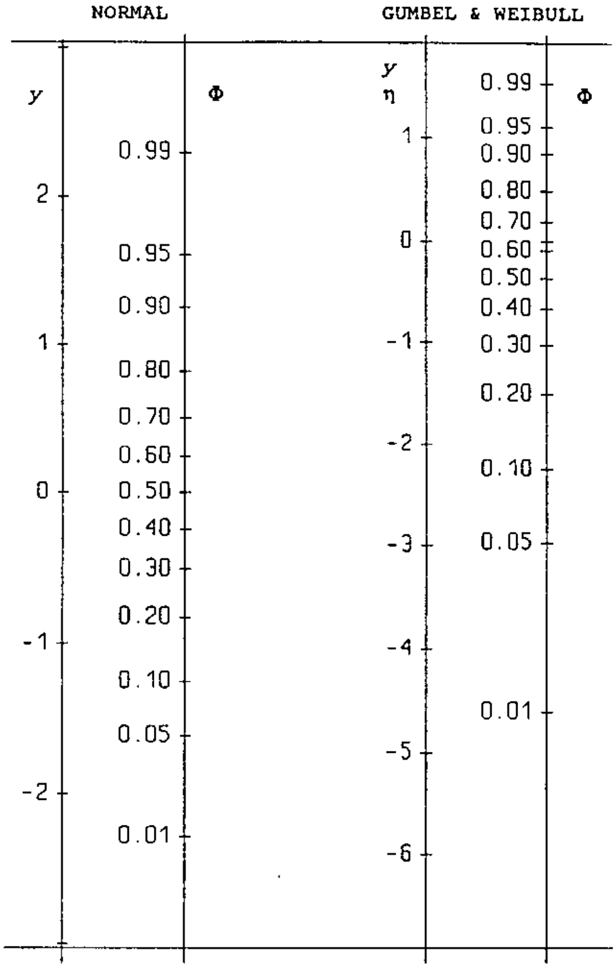


Fig.1 (A5) - Scales for the ordinates to be used in order to represent the Normal, the Gumbel and the Weibull distributions by straight lines. For Normal and Gumbel distributions the X-scale is linear. For the Weibull distribution it is logarithmic.

Often the scale of the standardized variable y is omitted and only the Φ scale is indicated on the probability paper.

From the classic expression of the Weibull distribution it is easy to obtain the relation:

$$y = \ln(\frac{1}{1 - \Phi}) \qquad \text{(A5.2)}$$

But, it can be seen from Table 4 that the standardized variable y is not a linear function of X. Taking the logarithm of y we obtain:

$$\ln(y) = \alpha\ln(X - X_o) - \alpha\ln(X_s) \qquad \text{(A5.3)}$$

From this expression it follows that in the special case of $X_0=0$ there is a linear relationship between lny and lnX. So for the new variables $\xi=\ln X$ and $\eta=\ln y$ we can apply the same procedure as described above for the Normal and Gumbel distributions. Taking into account (A5.2) we find:

$$\eta = \ln(y) = \ln[\ln(\frac{1}{1-\Phi})] \qquad (A5.4)$$

This means that the Φ scale to be used in this special case of the Weibull distribution is the same as that used for the Gumbel distribution. The difference between the two probability papers is in the X scale which is logarithmic for the Weibull and linear for the Gumbel paper.
Obviously, no straight line can be obtained when a Weibull distribution with $X_0\neq0$ is plotted on a Weibull probability paper.

APPENDIX 6

DETERMINATION OF THE PROBABILITY DISTRIBUTION OF THE RATIO "g"

As indicated in Section 8, the random variable g is defined by the ratio:

$$g = \frac{P_r}{P_{rmax}} \qquad (A6.1)$$

where:

P_r is the probability of obtaining the results, calculated with expression (8.1) for type F tests or (8.6) for type P tests, on the assumption that the failure probability is defined by the function $\Phi(X)$ that we have chosen (hypothesis H_0).

P_{rmax} is the maximum possible value of the probability of obtaining the results, calculated with expression (8.2), on the assumption that the failure probability associated with each interval of the variable X (for type F tests), or with each level of U (for type P tests), is equal exactly to the observed frequency of failure.

Wilks [8] has demonstrated that $-2\ln(g)$ tends asymptotically to the chi-square distribution i.e.:

$$-2\ln(g) = \chi^2(\nu) \qquad (A6.2)$$

where ν is the number of degrees of freedom determined as follows:

a) Type F tests:
ν = (Nr of intervals -1) - Nr of parameters

b) Type P tests:
ν = Nr of voltage levels - Nr of parameters

By means of expression (A6.2) the cumulative probability distribution of g can be obtained from that of χ^2. This is shown qualitatively in Fig.1(A6). In the upper part of the figure are plotted the diagrams of the chi-square cumulative distribution for different degrees of freedom ν. These cumulative curves indicate that the probability of χ^2 being less than a given value χ^2_o is P_o i.e.:

$$P(\chi^2\leq\chi^2_o) = P(\chi^2_o,\nu) = P_o \qquad (A6.3)$$

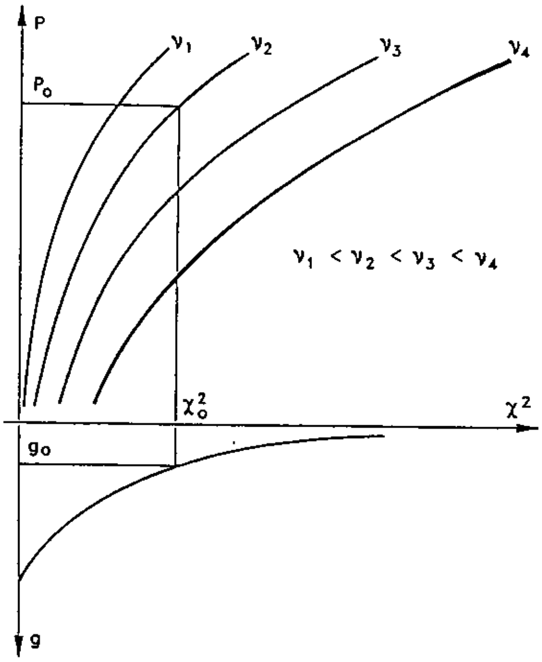


Fig.1 (A6) Diagrams of $P(\chi^2,\nu)$ and $g(\chi^2)$

In the lower part of Fig.1(A6) is plotted the curve $g = \exp[-\chi^2/2]$ as function of χ^2. It can be seen that to $\chi^2\leq\chi^2_o$ corresponds $g\geq g_o$ therefore :

$$P(g\leq g_o) = P(g_o) = 1 - P(\chi^2_o,\nu)$$

The last equation is valid for any value of g_o and represents the probability distribution of the ratio g. Therefore, the general expression of this distribution is:

$$P(g) = 1 - P(\chi^2,\nu) \qquad (A6.4)$$

The analytical expression of $P(\chi^2,\nu)$ and a partial representation of its tabulation are given at the end of this appendix.

From P(g) we can determine a limiting value g_{lim} for the ratio g such that:

$$P(g_{\text{lim}}) = P_{\text{lim}}$$

where P_{lim} is a sufficiently low, arbitrary value of the probability. Usual values for P_{lim} are 0.1, 0.05 and 0.01.

Then the hypothesis H_0 that the failure probability is given by $\Phi(X)$ is tested according to the following procedure:

- Selection of the value of P_{lim} and determination of the corresponding value of probability $P(\chi^2_{lim},\nu)$:

$$P(\chi^2_{\text{lim}}, \nu) = 1 - P_{\text{lim}} = C$$

- Determination of χ^2_{lim}, corresponding to the probability C and to the number of degrees of freedom ν, from tables of the χ^2 distribution [see Table 1(A6)]

- Determination of g_{lim} from the expression:

$$g_{\text{lim}} = \exp\left(-\frac{\chi^2_{\text{lim}}}{2}\right)$$

- Calculation of the ratio g from the experimental results by means of (A6.1).

- Rejection of H_0 if $g<g_{lim}$ or acceptance if $g \geq g_{lim}$.

THE CHI-SQUARE DISTRIBUTION

The analytical expression of the chi-square cumulative distribution function is defined as follows:

$$P(\chi^2, \nu) = \int_{\chi^2}^{\infty} f(w)\,dw$$

$$f(w) = \frac{w^{(\frac{\nu}{2}-1)}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}\exp(-\frac{w}{2})$$

$$\Gamma(y) = (y-1)! = \int_0^{\infty} t^{(y-1)}\exp(-t)\,dt$$

The tables of the chi-square distribution available from the relevant literature give the values of χ^2 as function of ν for a certain number of values of the probability. A partial reproduction of such table is given below.

TABLE 1 (A6) χ^2 as function of C and ν

ν	C=70%	C=80%	C=90%	C=95%	C=99%
1	1.07	1.64	2.71	3.84	6.63
2	2.41	3.22	4.61	5.99	9.21
3	3.67	4.64	6.25	7.81	11.3
4	4.88	5.99	7.78	9.49	13.3
5	6.06	7.29	9.24	11.1	15.1
6	7.23	8.56	10.6	12.6	16.8
7	8.38	9.80	12.0	14.1	18.5
8	9.52	11.0	13.4	15.5	20.1
9	10.7	12.2	14.7	16.9	21.7
10	11.8	13.4	16.0	18.3	23.2
11	12.9	14.6	17.3	19.7	24.7
12	14.0	15.8	18.5	21.0	26.2
13	15.1	17.0	19.8	22.4	27.7
14	16.2	18.2	21.1	23.7	29.1
15	17.3	19.3	22.3	25.0	30.6
16	18.4	20.5	23.5	26.3	32.0
17	19.5	21.6	24.8	27.6	33.4
18	20.6	22.8	26.0	28.9	34.8
19	21.7	23.9	27.2	30.1	36.2
20	22.8	25.0	28.4	31.4	37.6
21	23.9	26.2	29.6	32.7	38.9
22	24.9	27.3	30.8	33.9	40.3
23	26.0	28.4	32.0	35.2	41.6
24	27.1	29.6	33.2	36.4	43.0
25	28.2	30.7	34.4	37.7	44.3
26	29.2	31.8	35.6	38.9	45.6
27	30.3	32.9	36.7	40.1	47.0
28	31.4	34.0	37.9	41.3	48.3
29	32.5	35.1	39.1	42.6	49.6
30	33.5	36.3	40.3	43.8	50.9

Approximate expression of χ^2 for $\nu>30$:
$\chi^2_C = [Z_C + \sqrt{(2\nu-1)}]^2/2$ C=70% → $Z_C=Z_{70}=0.5244$
$Z_{80}=0.8416$ $Z_{90}=1.2816$ $Z_{95}=1.6449$ $Z_{99}=2.3263$

APPENDIX 7

EXAMPLES

Table of contents

EXAMPLE Nr 1

MULTIPLE LEVEL TEST

(Type P test)

TEST DATA

Test object: Rod-plane electrode configuration; 2 m air gap.

Test-voltage: Switching impulse of positive polarity; 400/4000 μs.

Test procedure:
- Different test voltages - multiple level method.
- Failure criterion: total breakdown of the gap.

Test conditions: Dry tests.

Experimental results:

U (kV_{peak})	N_f	N_W
795	0	20
815	2	18
835	0	20
850	4	16
870	7	13
890	11	9
910	12	8
930	19	1
950	18	2
970	20	0

STATISTICAL ANALYSIS

Interpolating function:
a) Modified Weibull; $K_o = 4$
b) Modified Weibull; $K_o = 2$

Likelihood function: Expression (6.6)

Most likely values of the parameters:
a) $U_{50M} = 889.7$ kV $Z_M = 42.3$ kV
b) $U_{50M} = 881.3$ kV $Z_M = 41.0$ kV

90% confidence intervals of the parameters:
a) $880.5 \text{ kV} \le U_{50} \le 898.8 \text{ kV}$
$34.5 \text{ kV} \le Z \le 53.2 \text{ kV}$
b) $871.9 \text{ kV} \le U_{50} \le 891.6 \text{ kV}$
$34.5 \text{ kV} \le Z \le 50.5 \text{ kV}$

Index of the goodness of fit:
a) P(g) = 0.238
b) P(g) = 0.0736

Diagrams: See Fig.1 (E1) and Fig.2 (E1)

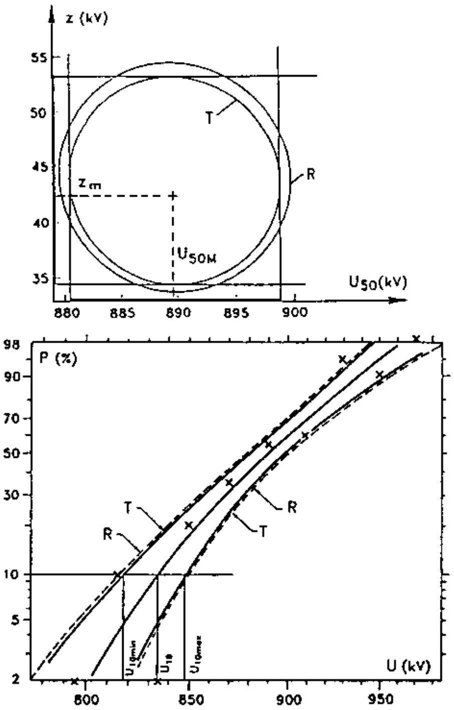


Fig.1 (E1) - Multiple level tests.
Case a) Acceptable goodness of fit i.e. P(g)>0.1. Diagrams of:
- *the locus R of the 90% confidence region of the parameters*
- *the locus T which extreme tangents determine the 90% confidence intervals of the parameters (see Section 9.3)*
- *the function P(U) with its 90% confidence bands R and T.*

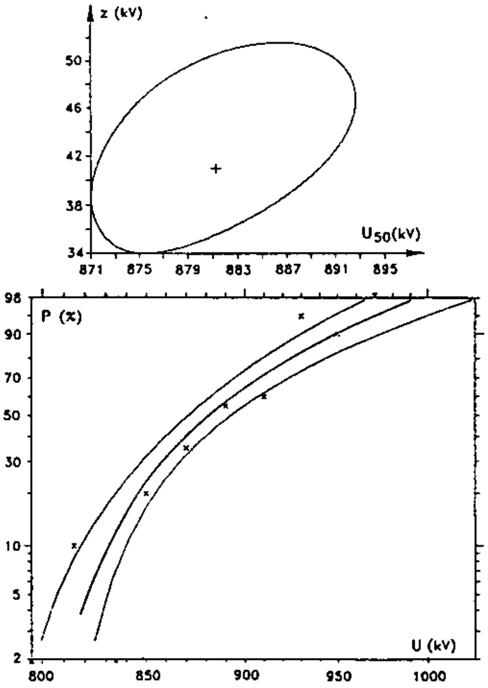


Fig.2 (E1) - Multiple level test.
Case b) Insufficient goodness of fit i.e. P(g)<0.1. Diagrams of:
- *the boundary of the 90% confidence region of the parameters (locus R)*
- *the function P(U) with its 90% confidence band (R).*

EXAMPLE Nr 2

UP AND DOWN TEST FOR THE DETERMINATION OF U_{50}

(type P test)

TEST DATA

Test object: Rod-rod electrode configuration; 4 m air gap.

Test-voltage: Switching impulse of positive polarity; 1800/7500 μs.

Test procedure:
- Different test-voltages; up and down method.
- Withstand procedure with $m_L = 1$.
- Failure criterion: total breakdown of the gap.

Test conditions: Dry tests.

Experimental results:

U (kV_{peak})	N_f	N_w
1770	0	2
1830	2	7
1890	7	8
1950	9	5
2010	5	0

STATISTICAL ANALYSIS

Interpolating function: Modif. Weibull; $K_0=4$

Likelihood function: Expression (6.6)

Most likely values of the parameters:
U_{50M} = 1904 kV Z_M = 86 kV

90% confidence intervals of the parameters:
1862 kV $\leq U_{50} \leq$ 1941 kV
52 kV $\leq Z \leq$ 205 kV

Index of the goodness of fit: P(g) = 0.663

Diagrams: See Fig.1 (E2)

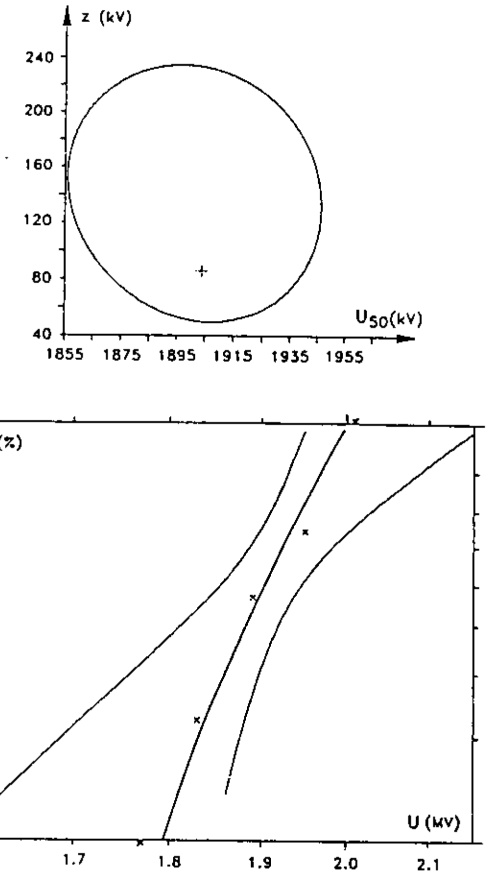


Fig.1 (E2) - Up and down tests. Diagrams of:
- *the boundary R of the 90% confidence region of the parameters*
- *the function P(U) with its 90% confidence band.*

EXAMPLE Nr 3

UP AND DOWN TEST (SHORT PROCEDURE) FOR THE DETERMINATION OF U_5

(type P test)

TEST DATA

Test object: Rod-rod electrode configuration; 4 m air gap.

Test-voltage: Switching impulse of positive polarity; 1800/7500 μs.

Test procedure:
- Different test voltages - up and down method.
- Withstand procedure with $m_L \leq 13$.
- Failure criterion: total breakdown of the gap.

Test conditions: Dry tests.

Experimental results:

U (kV_{peak})	N_f	N_w
1640	0	13
1670	2	29
1700	2	26
1730	2	35
1760	2	2

STATISTICAL ANALYSIS

Interpolating function: Modif. Weibull; $K_0=4$

Likelihood function: Expression (6.6)

Most likely values of the parameters:
U_{50M} = 1856 kV Z_M = 103 kV

90% confidence intervals of the parameters:
a) Without limitations about the "possible" values of the parameters: can not be determined.
b) With the limitation $Z \leq 400$ kV:
$1771 \text{ kV} \leq U_{50} \leq 2411 \text{ kV}$
$45 \text{ kV} \leq Z \leq 398 \text{ kV}$

50% confidence limits of the parameters:
Without limitations about the "possible" values of the parameters:
$1796 \text{ kV} \leq U_{50} \leq 2130 \text{ kV}$
$62 \text{ kV} \leq Z \leq 287 \text{ kV}$

Index of the goodness of fit: P(g) = 0.192

Diagrams: See Fig.1 (E3)

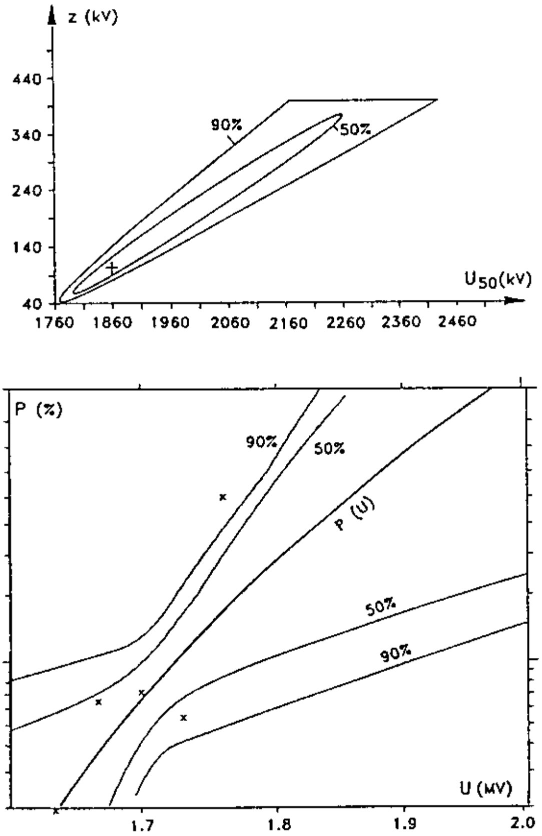


Fig.1 (E3) - Up and down tests with $m_L \leq 13$. Diagrams of:
- *the boundary of the 90% confidence region of the parameters with the condition: $Z \leq 400$ kV*
- *the boundary of the 50% confidence region without conditions*
- *the function P(U) with its 90% and 50% confidence bands.*

EXAMPLE Nr 4

UP AND DOWN TEST (COMPLETE PROCEDURE) FOR THE DETERMINATION OF U_5

(type P test)

TEST DATA

Test object: Rod-rod electrode configuration; 4 m air gap.

Test-voltage: Switching impulse of positive polarity; 1800/7500 µs.

Test procedure:
- Different test voltages - up and down method.
- Withstand procedure with $m_L = 13$.
- Failure criterion: total breakdown of the gap.

Test conditions: Dry tests.

Experimental results:

U (kV_{peak})	N_f	N_w
1640	0	13
1670	2	37
1700	3	36
1730	4	48
1760	6	20

STATISTICAL ANALYSIS

Interpolating function: Modif. Weibull; $K_0=4$

Likelihood function: Expression (6.6)

Most likely values of the parameters:
U_{50M} = 1844 kV Z_M = 95 kV

90% confidence intervals of the parameters:
$1785 \text{ kV} \leq U_{50} \leq 2356 \text{ kV}$
$52 \text{ kV} \leq Z \leq 471 \text{ kV}$

Index of the goodness of fit: P(g) = 0.585

Diagrams: See Fig.1 (E4)

N.B. The results of the test in Example Nr 3 are obtained from those of this example considering, for each voltage level, only the applications until the first discharge (including it).

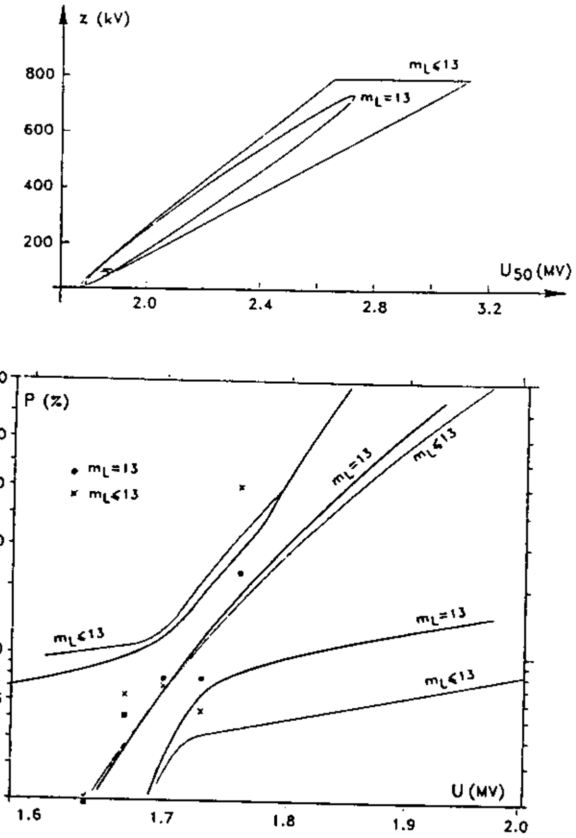


Fig.1 (E4) - Comparison of up and down tests performed with the complete and with the short procedures. Diagrams of:

- *the boundaries of the 90% confidence regions of the parameters*
- *the functions P(U) with their 90% confidence band.*

EXAMPLE Nr 5

RAMP TEST (type F test)

TEST DATA

Test object: Point-to-sphere 50mm gap in mineral oil

Test-voltage: AC (50Hz) ramp voltage with rate of rise r = 1 kV/s

Test procedure:

- Voltage increased until failure. Circuit opened immediately after failure. One minute interval between failure and successive test voltage application.
- Failure criterion: appearance of a PD pulse with apparent charge exceeding 100 pC

Test conditions: Test cell at atmospheric pressure and ambient temperature

Experimental results:

20 values of the PD inception voltage U (kV) given in chronological order (moving horizontally)

23.9	22.8	23.4	25.0	22.8
21.7	22.2	23.3	22.8	22.8
24.3	20.5	23.8	23.5	23.5
25.2	22.4	24.6	23.4	22.3

N.B. The arrangement of the input data for the calculations is shown below. The results reported in Table 3 are taken from this example.

INPUT DATA ARRANGEMENT FOR THE CALCULATIONS

i	U_i (kV)	N_i	ΣN_i	U^*_i (kV)	$F_e(U_i)$ (%)
1	20.5	1	1	21.1	5.0
2	21.7	1	2	22.0	10.0
3	22.2	1	3	22.3	15.0
4	22.3	1	4	22.4	20.0
5	22.4	1	5	22.6	25.0
6	22.8	4	9	23.0	45.0
7	23.3	1	10	23.4	50.0
8	23.4	2	12	23.5	60.0
9	23.5	2	14	23.6	70.0
10	23.8	1	15	23.9	75.0
11	23.9	1	16	24.1	80.0
12	24.3	1	17	24.5	85.0
13	24.6	1	18	24.8	90.0
14	25.0	1	19	25.1	95.0
15	25.2	1	20	∞	100.0

i - number of the voltage level
U_i - value of the failure voltage
N_i - frequency of failures at the voltage level U_i

$\Sigma N_i = \sum_{k=1}^{i} N_k$ - cumulative frequency of failures up to U_i

$U^*_i = (U_{i+1} + U_i)/2$ - upper limit of the interval "i"

$F_e(U_i)$ - estimated value of $F(U_i)$

N.B. In the calculations with the likelihood function given by expression (6.9) that is:

$$L = \prod_{i=1}^{m} \varphi(X_i)^{N_i}$$

the values of X_i are: $X_i=U_i$. In the calculations with the likelihood function given by expression (6.11) that is:

$$L = \prod_{i=1}^{m} \Delta\Phi_i^{N_i} = \prod_{i=1}^{m} [\Phi(x_i) - \Phi(x_{i-1})]^{N_i}$$

the values of x_i are: $x_i = U^*_i$.

STATISTICAL ANALYSIS

Interpolating function: Modif. Weibull; K_0=4

Likelihood function:
a) Expression (6.9)
b) Expression (6.11)

Most likely values of the parameters:
a) U_{50M} = 23.26 kV Z_M = 1.18 kV
b) U_{50M} = 23.28 kV Z_M = 1.22 kV

90% confidence intervals of the parameters:
a) $22.74 \text{ kV} \le U_{50} \le 23.75 \text{ kV}$
$0.92 \text{ kV} \le Z \le 1.64 \text{ kV}$
b) $22.73 \text{ kV} \le U_{50} \le 23.80 \text{ kV}$
$0.91 \text{ kV} \le Z \le 1.77 \text{ kV}$

Index of the goodness of fit:
a) Not calculated
b) P(g) = 0.952

Diagrams: See Fig.1 (E5)

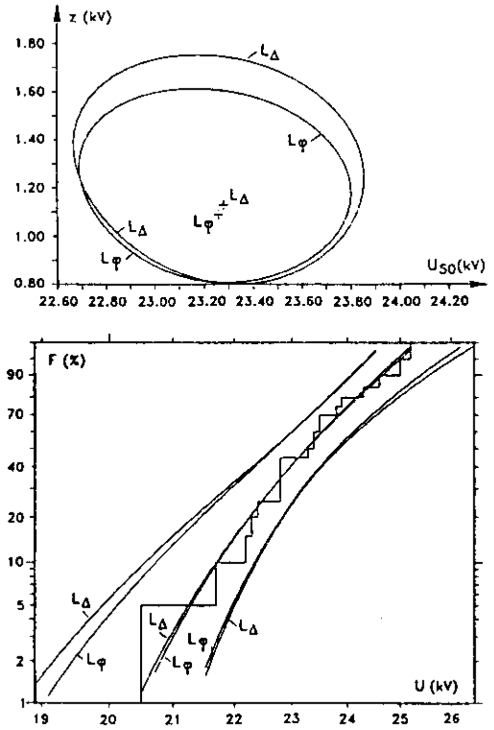


Fig.1 (E5) - Ramp tests. Comparison of calculations made with the likelihood functions (6.9) indicated with L_φ and (6.12) indicated with L_Δ. Diagrams of:
- *the boundaries of the 90% confidence regions of the parameters*
- *the functions F(U) with their 90% confidence bands.*

EXAMPLE Nr 6

STAIR TEST (type F test)

TEST DATA

Test object: Cylinder-to-plane gap in transformer oil. Crome plated cylinder (diameter 25 mm). Gap size: 50 mm.

Test-voltage: AC stair voltage with U_1 = 150 kV ΔU = 10 kV; t_s = 1 min.

Test procedure:
- Voltage increased until failure. Circuit opened immediately after failure. 10 minutes interval between failure and successive test voltage application.
- Failure criterion: complete breakdown of the gap

Test conditions: Test tank at atmospheric pressure and ambient temperature. Cylinder in vertical position. Change of oil after each 20 breakdowns.

Experimental results:

100 values of the breakdown voltage U (kV)
given in chronological order
(moving horizontally)

340	380	310	260	290
300	320	350	350	370
310	360	350	280	290
310	290	290	270	360
210	340	370	270	320
350	270	270	340	330
300	290	300	340	280
350	350	340	340	360
330	250	300	300	320
250	300	290	350	360
280	380	240	300	340
250	330	350	380	360
300	420	380	380	330
270	270	330	320	350
350	260	310	340	280
290	310	310	360	340
320	270	340	380	340
390	260	310	360	330
270	290	320	250	310
240	260	330	280	380

INPUT DATA ARRANGEMENT FOR THE CALCULATIONS

i	U_i (kV)	N_i	ΣN_i
1	210	1	1
2	240	2	3
3	250	4	7
4	260	4	11
5	270	8	19
6	280	5	24
7	290	8	32
8	300	8	40
9	310	8	48
10	320	6	54
11	330	7	61
12	340	11	72
13	350	10	82
14	360	7	89
15	370	2	91
16	380	7	98
17	390	1	99
18	420	1	100

N.B. In contrast with what indicated for the ramp tests, in this case the variable is: $x_i = X_i = U_i$

STATISTICAL ANALYSIS

A. ANALYSIS OF THE FIRST 20 RESULTS

Interpolating function: Modif. Weibull; $K_o=3$

Likelihood function: Expression (6.11)

Most likely values of the parameters:
U_{50M} = 313.7 kV Z_M = 36.0 kV

90% confidence intervals of the parameters:
297.0 kV ≤ U_{50} ≤ 329.8 kV
27.4 kV ≤ Z ≤ 51.1 kV

Index of the goodness of fit: P(g) = 0.398

Diagrams: See Fig.1(E6) and Fig.2(E6)

B. ANALYSIS OF ALL 100 RESULTS

Interpolating function: Mod. Weibull; $K_o=3.5$

Likelihood function: Expression (6.11)

Most likely values of the parameters:
U_{50M} = 312.6 kV Z_M = 43.0 kV

90% confidence intervals of the parameters:
304.5 kV ≤ U_{50} ≤ 320.5 kV
38.0 kV ≤ Z ≤ 49.4 kV

Index of the goodness of fit: P(g) = 0.573

Diagrams: See Fig.1(E6) and Fig.2(E6).

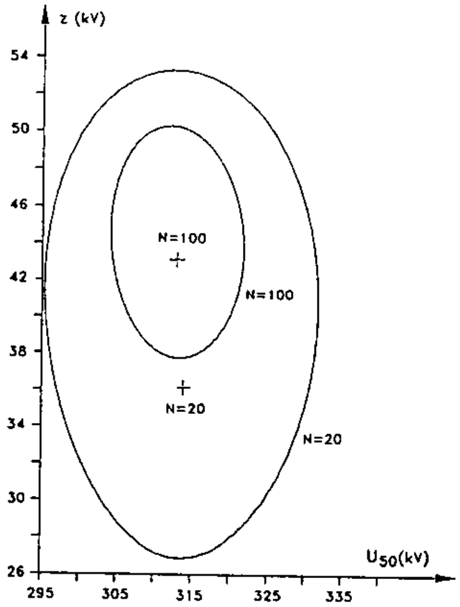


Fig.1 (E6) - Stair tests. Diagrams of the boundaries of the 90% confidence regions of the parameters for the cases of N=20 and N=100 results.

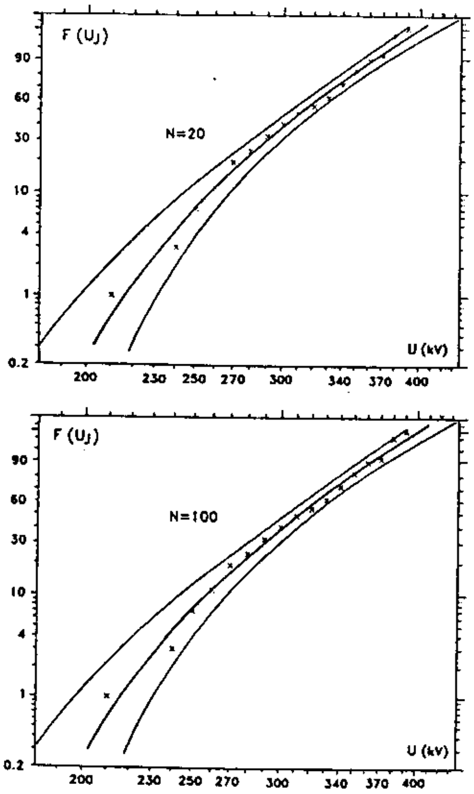


Fig.2 (E6) - Stair tests. Conventional representation with continuous lines of the functions $F(U_j)$ and their 90% confidence bands for N=20 and N=100 results.

EXAMPLE Nr 7

IMPULSE TESTS PERFORMED WITH THE STEP METHOD

TEST DATA

Test object: Point-to-sphere 15 mm gap in mineral oil.

Test-voltage: Lightning impulse 1.2/50 μs positive polarity

Test procedure: Application of impulses with progressively increasing peak values (step method) - one impulse per level until breakdown. Initial voltage level: U_1=20 kV. Voltage step: ΔU= 5 kV.

Test conditions: Test cell at atmospheric pressure and ambient temperature. Electrodes and oil changed after each breakdown.

Experimental results:

10 values of the breakdown voltage (kV peak) in chronological order (moving horizontally)

75	60	55	60	75
60	55	50	45	60

INPUT DATA ARRANGEMENT FOR THE CALCULATIONS

A. Considering the data as obtained from the application of different test-voltages (test-voltage = single impulse)

i	U (kV_{pk})	N_f	N_w
1	45	1	9
2	50	1	8
3	55	2	6
4	60	4	2
5	75	2	0

B. Considering the data as obtained from the application of complex stair test-voltages (test-voltage = series of incr. impulses)

i	U_i (kV_{pk})	N_i	ΣN_i
1	45	1	1
2	50	1	2
3	55	2	4
4	60	4	8
5	75	2	10

STATISTICAL ANALYSIS

A. TEST-VOLTAGE = SINGLE IMPULSE. ESTIMATION OF P(U)

Interpolating function: Modif. Weibull; K_o=4

Likelihood function: Expression (6.6)

Most likely values of the parameters:
U_{50M} = 58.4 kV Z_M = 8.31 kV

90% confidence intervals of the parameters:
54.4 kV ≤ U_{50} ≤ 68.1 kV
4.66 kV ≤ Z ≤ 21.8 kV

Index of the goodness of fit: P(g) = 0.811

Diagrams: See Fig.1(E7) and Fig.2(E7)

B. TEST-VOLTAGE = COMPLEX STAIR (SERIES OF INCREASING IMPULSES). ESTIMATION OF F(U_j).

Interpolating function: Modif. Weibull; K_o=4

Likelihood function: Expression (6.11)

Most likely values of the parameters:
U_{50M} = 55.4 kV Z_M = 6.88 kV

90% confidence intervals of the parameters:
50.4 kV ≤ U_{50} ≤ 60.3 kV
4.30 kV ≤ Z ≤ 13.1 kV

Index of the goodness of fit: P(g) = 0.835

Diagrams: See Fig.1(E7) and Fig.2(E7)

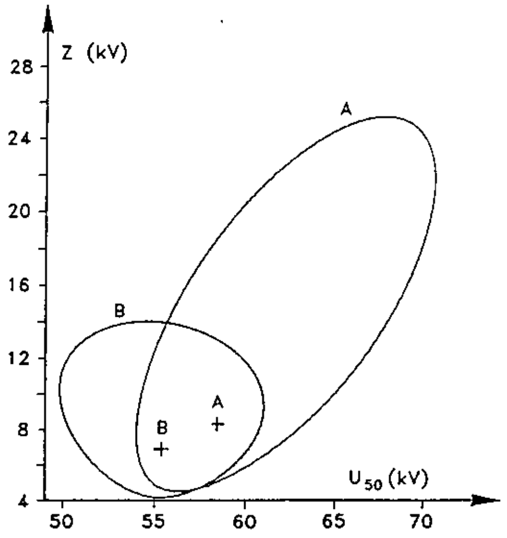


Fig.1 (E7) - 90% confidence regions of the parameters of the functions:
P(U) - case A (test-voltage = single imp.)
F(U_j) - case B (test-voltage = series of increasing impulses).

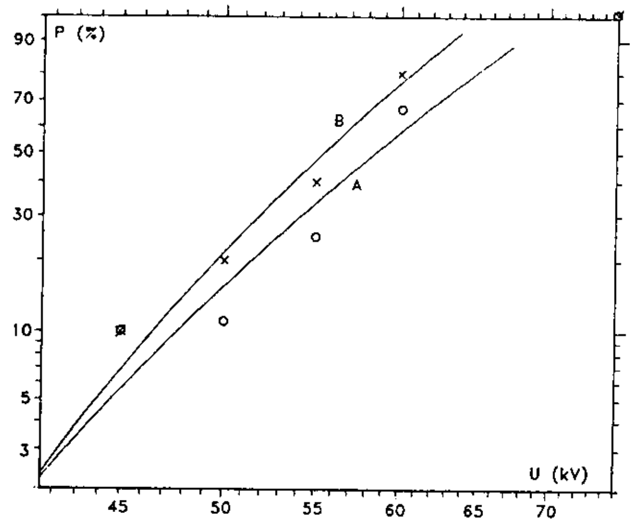


Fig.2 (E7) - Conventional representation, with continuous curves, of the discrete functions P(U) and F(U_j) estimated from the same set of results, analysed in two ways (cases A and B).

EXAMPLE Nr 8

COMPARISON TEST

TEST DATA

Test object 1: Point-to-sphere 50mm gap in mineral oil. Extension of the needle from the tip of the needle holder equal to 20 mm.

Test object 2: The same as test object 1 but the extension of the needle is equal to 15 mm.

Test voltage: AC (50Hz) ramp voltage with rate of rise r = 1 kV/s

Test procedure:
- Voltage increased until failure. Circuit opened immediately after failure. One minute interval between failure and successive test voltage application.
- Failure criterion: appearance of a PD pulse with apparent charge exceeding 100 pC

Test conditions: Test cell at atmospheric pressure and ambient temperature

Experimental results:

20 PD inception voltages (kV) for each test object given below in chronological order (reading horizontally)

Test object 1 (results R_1)

19.9	21.5	23.3	23.3	22.2
23.3	23.3	24.4	24.4	22.0
22.1	23.2	24.1	23.5	24.2
23.6	23.6	24.7	24.1	24.3

Test object 2 (results R_2)

21.2	23.5	23.5	24.7	23.0
24.8	24.2	27.0	24.7	23.6
23.6	25.2	25.3	22.5	24.2
24.1	24.3	24.8	23.7	26.0

STATISTICAL ANALYSIS

The purpose of the experiment is to check if the extension of the needle from the tip of the needle holder has an effect on the PD inception voltage level. For this reason a statistical test of the following hypotheses is made:

H_0: The two sets of results come from the same population. In other words the extension of the needle has no effect on the PD inception voltage and therefore the differences between the results of the two sets are not significant.

H_1: The opposite of H_0 i.e. the two sets of results do not come from the same population. The extension of the needle has an effect on the PD inception voltages and therefore the differences between the results of the two sets are significant.

Test based on the statistic b

For the determination of the value of the ratio b by expression (10.7) it is necessary to determine the intervals on the X-axis (X=U). This determination is not completely arbitrary and should be done according to the following criteria:

- The intervals must cover the whole range of voltages from 0 to ∞.
- Any interval should contain at least one result of each set.
- The limits of the intervals should not coincide with results.
- The maximum number of intervals, satisfying the above requirements, should be selected (unless the results are extremely numerous).

In order to facilitate the determination of the intervals, the results of the two sets are arranged in increasing order as shown below. The two sets can be distinguished because the number of results obtained from test object 1 at a given voltage level are marked with an asterix.

U (kV)	N_i
19.9	1^*
21.2	1
21.5	1^*
22.0	1^*
22.1	1^*
22.2	1^*
22.5	1
23.0	1
23.2	1^*
23.3	4^*
23.5	1^*; 2
23.6	2^*; 2

U (kV)	N_i
23.7	1
24.1	2^*; 1
24.2	1^*; 2
24.3	1^*; 1
24.4	2^*
24.7	1^*; 2
24.8	2
25.2	1
25.3	1
26.0	1
27.0	1
-	-

The intervals obtained from these tables with the application of the criteria mentioned above are given in the table below together with the data necessary for the calculation of the ratio "b".

INTERVAL	N_{1i}	N_{1i}/N_1	N_{2i}	N_{2i}/N_2	N_i	N_i/N
0.0 - 21.4	1	0.05	1	0.05	2	0.050
21.4 - 22.6	4	0.20	1	0.05	5	0.125
22.6 - 23.4	5	0.25	1	0.05	6	0.150
23.4 - 23.9	3	0.15	5	0.25	8	0.200
23.9 - 24.6	6	0.30	4	0.20	10	0.250
24.6 - inf.	1	0.05	8	0.40	9	0.225

$$b = \frac{\prod_{i=1}^{6}\left(\frac{N_i}{N}\right)^{N_i}}{\prod_{i=1}^{6}\left(\frac{N_{1i}}{N_1}\right)^{N_{1i}}\left(\frac{N_{2i}}{N_2}\right)^{N_{2i}}} = 2.55\times10^{-3}$$

$$\nu = 6-1 = 5 \qquad \chi^2{}_{lim} = \chi^2(5,90) = 9.24$$
$$b_{lim} = \exp(-9.24/2) = 9.85\cdot10^{-3}$$

$$\mathbf{b < b_{lim} \rightarrow H_o \text{ is rejected}}$$

Test based on the verification of the goodness of fit

Interpolating curve relative to the set of results R_1

Interpolating function: Modif. Weibull; K_o=4

Likelihood function: Expression (6.11)

Most likely values of the parameters:
U_{50M} = 23.3 kV Z_M = 1.19 kV

90% confidence intervals of the parameters:
22.78 kV $\leq U_{50} \leq$ 23.80 kV
0.90 kV $\leq Z \leq$ 1.72 kV

Index of the goodness of fit:
P(g) = 0.264

Interpolating curve relative to the set of results R_2

Interpolating function: Modif. Weibull; K_o=4

Likelihood function: Expression (6.11)

Most likely values of the parameters:
U_{50M} = 24.2 kV Z_M = 1.34 kV

90% confidence intervals of the parameters:
23.64 kV $\leq U_{50} \leq$ 24.81 kV
1.00 kV $\leq Z \leq$ 1.94 kV

Index of the goodness of fit: P(g) = 0.796

Interpolating curve relative to the two sets of results R_1 and R_2

Interpolating function: Modif. Weibull; K_o=4

Likelihood function: Expression (6.11)

Most likely values of the parameters:
U_{50M} = 23.8 kV Z_M = 1.38 kV

90% confidence intervals of the parameters:
22.74 kV $\leq U_{50} \leq$ 23.75 kV
0.92 kV $\leq Z \leq$ 1.64 kV

Index of the goodness of fit: P(g) = 0.443

From the above analysis it follows that the expression of the interpolating function is:

$$\Phi = 1 - 0.5^{y}$$
$$y = \left[1 + \frac{U - 23.8}{5.52}\right]^{4.83} \qquad \text{(E8.1)}$$

Verification of the goodness of fit of (E8.1) to the results R_1

This verification is based on the ratio g which in this case is given by the expression:

$$g = \frac{\prod_{i=1}^{m}\Delta\Phi_i{}^{N_{1i}}}{\prod_{i=1}^{m}\left(\frac{N_{1i}}{N_1}\right)^{N_{1i}}} \qquad \text{(E8.2)}$$

where $\Delta\Phi_i=\Phi(U_i)-\Phi(U_{i-1})$.

The data relative to the set of results R_1 necessary for the determination of g are given in the table below.

Results R_1

i	INTERVAL	N_i	N_i/N	Φ_i	$\Delta\Phi_i$
1	0.0 - 20.7	1	0.05	0.0128	0.0128
2	20.7 - 21.8	1	0.05	0.0759	0.0631
3	21.8 - 22.1	1	0.05	0.1105	0.0346
4	22.1 - 22.2	1	0.05	0.1243	0.0138
5	22.2 - 22.7	1	0.05	0.2110	0.0867
6	22.7 - 23.3	1	0.05	0.3548	0.1438
7	23.3 - 23.4	4	0.20	0.3824	0.0276
8	23.4 - 23.6	1	0.05	0.4401	0.0577
9	23.6 - 23.9	2	0.10	0.5304	0.0903
10	23.9 - 24.2	2	0.10	0.6216	0.0912
11	24.2 - 24.3	1	0.05	0.6513	0.0297
12	24.3 - 24.4	1	0.05	0.6805	0.0292
13	24.4 - 24.6	2	0.10	0.7362	0.0512
14	24.6 - inf.	1	0.05	1.0000	0.2638

Applying expression (E8.2) to the data of the above table we obtain:

$$g = 4.19 \cdot 10^{-5} \qquad \nu = 14 - 1 = 13$$

$$\chi^2_{lim} = \chi^2(13,90) = 19.8$$

$$g_{lim} = \exp(-19.8/2) = 6.77 \cdot 10^{-5}$$

$g < g_{lim}$ → H_0 **is rejected**

Verification of the goodness of fit of (E8.1) to the results R_2

This verification is based on the ratio g which in this case is given by the expression:

$$g = \frac{\prod_{i=1}^{m} \Delta\Phi_i^{N_{2i}}}{\prod_{i=1}^{m} \left(\frac{N_{2i}}{N_2}\right)^{N_{2i}}} \qquad \text{(E8.3)}$$

where $\Delta\Phi_i = \Phi(U_i) - \Phi(U_{i-1})$.

The table with the data relative to the results R_2 necessary for the calculation of g is given below.

Results R_2

i	INTERVAL	N_i	N_i/N	Φ_i	$\Delta\Phi_i$
1	0.0 - 21.9	1	0.05	0.0864	0.0864
2	21.9 - 22.8	1	0.05	0.2320	0.1456
3	22.8 - 23.3	1	0.05	0.3548	0.1228
4	23.3 - 23.6	2	0.10	0.4401	0.0853
5	23.6 - 23.7	2	0.10	0.4698	0.0297
6	23.7 - 23.9	1	0.05	0.5304	0.0606
7	23.9 - 24.2	1	0.05	0.6216	0.0912
8	24.2 - 24.3	2	0.10	0.6513	0.0297
9	24.3 - 24.5	1	0.05	0.7088	0.0575
10	24.5 - 24.8	2	0.10	0.7876	0.0788
11	24.8 - 25.0	2	0.10	0.8335	0.0459
12	25.0 - 25.3	1	0.05	0.8907	0.0572
13	25.3 - 25.6	1	0.05	0.9334	0.0427
14	25.6 - 26.5	1	0.05	0.9913	0.0579
15	26.5 - inf.	1	0.05	1.0000	0.0087

Applying expression (E8.3) to the data of the above table we obtain:

$$g = 4.58 \cdot 10^{-3} \qquad \nu = 15 - 1 = 14$$

$$\chi^2_{lim} = \chi^2(14,90) = 21.1$$

$$g_{lim} = \exp(-21.1/2) = 2.62 \cdot 10^{-5}$$

$g > g_{lim}$ → H_0 **can not be rejected**

Final conclusion

For the rejection of the hypothesis H_0 it is sufficient that we find $g < g_{lim}$ for at least one of the sets of results. Since this is the case for the results R_1 the final conclusion is:

the hypothesis H_0 is rejected

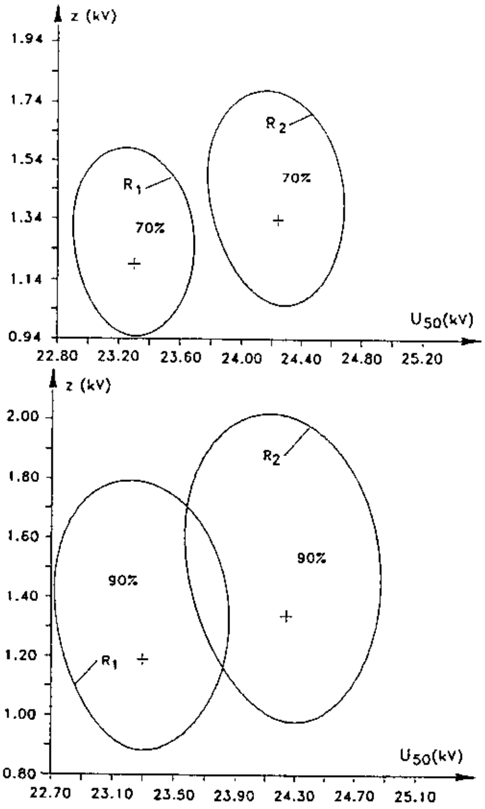


Fig.1 (E8) - Comparison tests. Diagrams of the boundaries of the 90% and 70% confidence regions of the parameters relative to the sets of results R_1 and R_2.

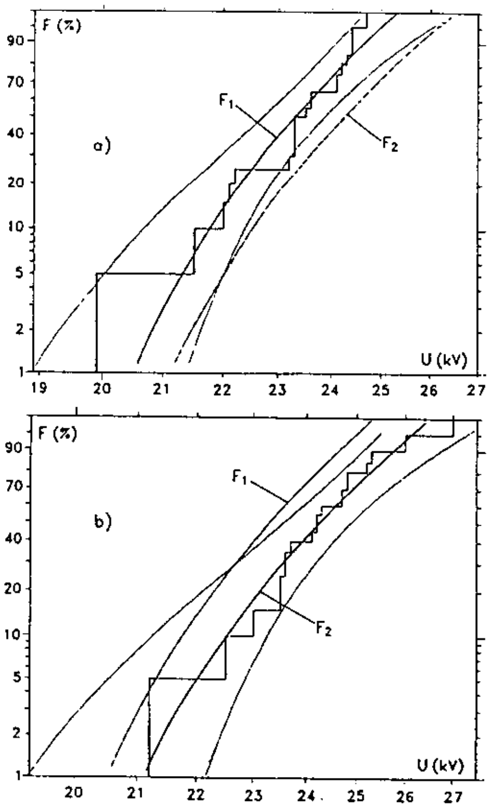


Fig.2 (E8) - Comparison tests. Diagrams of:

a) the function $F_1(U)$ with its 90% confidence band compared with the function $F_2(U)$ plotted with dotted line

b) the function $F_2(U)$ with its 90% confidence band compared with the function $F_1(U)$ plotted with dotted line.

APPENDIX 8

EXTRACT FROM REFERENCE [1]

STATISTICAL EVALUATION OF DIELECTRIC TEST RESULTS

by G. Carrara and W. Hauschild
Electra number 133, December 1990

Reference [1] covers the following aspects.

a) The description and the classification of the tests considered in IEC Publication 60 [5]:
  Class-1: Multiple Level Method (MLM),
  Class-2: Up-and-Down Method (UDM),
  Class-3: Progressive Stress Method (PSM). covering all procedures indicated as "Equal Test-voltages" in Table 1.

b) The concept of "Reaction function" to characterize the response of the insulation to the application of a test-voltage. Therefore, all functions in Table 1 of Section 4 are the reaction functions to the listed test-voltages. When the "interpolating function" mentioned in Section 4 is an estimate of a function belonging to Table 1, it is the estimate of the relevant reaction function.

c) The mathematical expressions (Gauss, Weibull and Gumbel distributions) used for the approximations of the characteristic functions.

d) The analysis of the procedures for the best point estimation of the parameters, for the evaluation of their confidence, and for the graphical representation of the empirical and of the estimated characteristic functions using the "classical" and the "likelihood" methods.

In this extract only the parts relevant to the "reaction function" and to the analysis with the "classical" methods are reported.

To apply the statistical analysis it is assumed that the insulation reacts to the application of a test voltage with a discharge probability, which depends on the characteristics of the test voltage. The set of all test voltages with the same shape, which are characterized by their peak value, will give discharge probabilities ranging from zero to one. The relationship between the probability of discharge P, and the corresponding peak value of the test voltage U, is a functional relationship, P = P(U), that can be called "Reaction function" of the insulation to that particular test voltage shape.

For most practical types of insulation and of test voltage shapes, the Reaction function shows a monotonous increase of the discharge probability with the increase of the peak value of the test voltage (Fig. 1 (A8) curve a). However, in some types of insulation, the discharge processes are sensitive to the absolute value of some characteristics of the test voltage (absolute steepness of the front of impulse voltages, etc.), which, with constant shape, vary with the peak value of the test voltage. In such cases, more than one discharge mechanism is effective, and, in a particular voltage range, the discharge probability may even be decreasing with the voltage (Fig. 1 (A8) b).

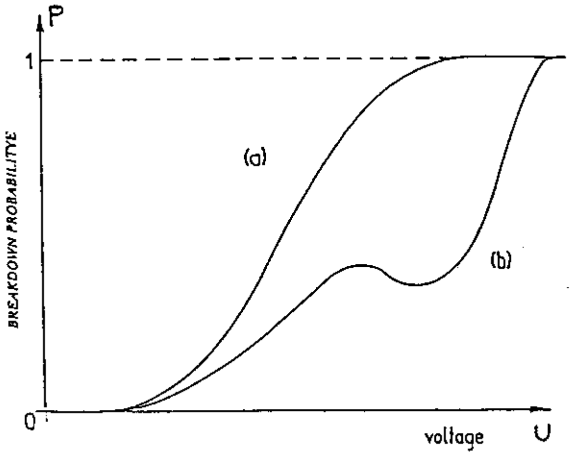


Fig. 1 (A8) - Shape of reaction function for impulse voltages.
a) Commonly encountered shape: P is monotonously increasing with the peak value U of u(t).
b) Shape for a particular dependence of the breakdown mechanism on the steepness of the front of u(t).

Tests performed with MLM and UDM estimate directly the reaction function. The MLM estimates the values of the parameters which characterize the whole function. The UDM gives estimates of voltages U_p with desired target probabilities p'.

The reaction function P(U) can be estimated also using the PSM. This occurs when the PSM consists of the application of series of test voltages with increasing peak values until the first discharge occurs. To estimate the parameter of the reaction function, the discharge probability should be derived from the frequency of discharge for each level separately.

It should be recognized that the reaction function, P(U), and the cumulative probability function of the discharge voltages, F(U), are theoretically different. Actually, F(U), that is the probability of having a discharge before or at the level U, depends not only on the discharge probability at level U, but also on the fact that all levels at lower peak voltages were withstood. Furthermore, F(U) depends also on the test procedure (starting voltage U and step amplitude ΔU). A relation, however, exists between F(U) and P(U) for each test procedure.

The examination of this consideration led to the introduction of the term "complex test-voltage" in Section 2.1, necessary to clarify the matter. Actually, the function P(U) considers the impulses as ***single test voltages****, and the stair as a method of their application, while F(U) considers the series of impulses as a* ***complex test-voltage****, which includes starting voltage U and step amplitude ΔU in its characterization.*

Reference [1] then describes the "classical methods" for the analysis of test results, preceded by a description of the "Procedure for the analysis".

The analysis of test results consists of:

An independence test, to check if the results show a trend (e.g. higher discharge frequency at the beginning than at the end of a series), or too many or too few groups of consecutive iterations of the same event.

A graphical representation of the empirical distribution function (i.e. the function as measured).

A mathematical approximation of this function by a suitable expression, using the procedure described for the determination of the parameters of the best fitting expression.

A goodness-of-fit test, to check if the best fitting mathematical approximation fits acceptably the test results.

All statistic tests, in particular the confidence estimations of the parameters of the expressions, are made on the basis of a preselected statistic confidence C (typical values of C are 0.9 to 0.99). In case of mixed distributions, which are observed relatively often, partial approximations by the mentioned mathematical expressions, or special methods described in the literature, are recommended.

Analysis of Class-1 tests (MLM)

The independence test. An example is given in Table 1 (A8). Firstly a trend test is performed. It simply consists of subdividing the whole sample in two parts, and evaluating the test statistic z_t:

$$z_t = \frac{|f_1 - f_2|}{\sqrt{A*B}} \qquad \text{(A8.1)}$$

$A = (f_1*n_1 + f_2*n_2)/(n_1 + n_2)$
$B = (n_1*(1 - f_1) + n_2*(1 - f_2))/(n_1*n_2)$
f_1 and f_2 are the discharge frequency,
n_1 and n_2 the sample size.

If the statistic z_t is not larger than the critical value of the Gauss distribution $\lambda_{(1+C)/2}$, corresponding to a desired confidence (e.g. C= 0.95, $\lambda_{0.975}$ = 1.96), the independence is not rejected.

Table 1 (A8) - Independence test for the discharge probability.

i	Events : Discharge x Withstand										f_i	f_n
1 - 10	-	-	x	-	-	-	x	x	-	-	0.30	0.29
11 - 20	-	x	-	-	x	-	-	-	-	-	0.20	
21 - 30	x	-	-	x	-	-	x	-	x	-	0.40	
31 - 40	-	-	-	-	x	-	-	x	-	-	0.20	
41 - 50	-	-	-	-	x	-	x	x	-	-	0.30	
51 - 60	-	-	x	-	-	x	-	-	x	-	0.30	
61 - 70	-	-	-	x	-	-	-	x	-	-	0.20	
71 - 80	x	x	x	-	-	x	-	-	x	-	0.50	
81 - 90	-	-	x	-	-	-	-	-	-	-	0.10	
91 - 100	x	-	-	-	x	x	-	-	-	x	0.40	
iteration test	Sample size n = 100 Discharges k = 29 Iterations r = 48						$\lambda_{0.975}$ = 1.960 z_i = 1.656 < $\lambda_{0.975}$					
probability trend test	f_{1-10} = 0.30				f_{91-100} = 0.40				z_{t1} = 0.469 < $\lambda_{0.975}$			
	f_n = 0.29				$f_{i,min}$ = 0.10				z_{tm} = 1.286 < $\lambda_{0.975}$			
	f_n = 0.29				$f_{i,max}$ = 0.50				z_{tx} = 1.370 < $\lambda_{0.975}$			

A runs test of iterations is then performed. It consists of subdividing the whole sample of size n in sequences of the same event, counting their number, r, and evaluating the test statistic z_i:

$$z_i = \frac{r - 2(n-k)\dfrac{k}{n}}{2(n-k)\dfrac{k}{n^{1.5}}} \qquad \text{(A8.2)}$$

If the statistic z_i is not larger than the critical value of the Gauss distribution, $\lambda_{(1+C)/2}$, corresponding to a desired confidence (e.g. C = 0.95, $\lambda_{0.975}$ = 1.96), the independence is not rejected.

It should be mentioned that the two statistics z_t and z_i approach Gaussian distribution in large samples; for small samples, critical values for these statistics are available in the relevant statistical tables.

The graphical representation. For each value x_j the corresponding discharge frequency f_j and its C-confidence interval $P_{li} \longleftrightarrow P_{ui}$ are plotted on the probability paper of the approximating function (Fig. 2 (A8)). Then a by-eye best fitting straight line is drawn through the confidence intervals, representing empirically the approximating function. This can be used for the estimation of parameters by quantiles (Fig. 2 (A8)). It is, however, much better to draw the straight line according to the results of an analytical evaluation.

The analytical evaluation. The measured frequencies f_j are transformed, using the inverse of the mathematical expression chosen

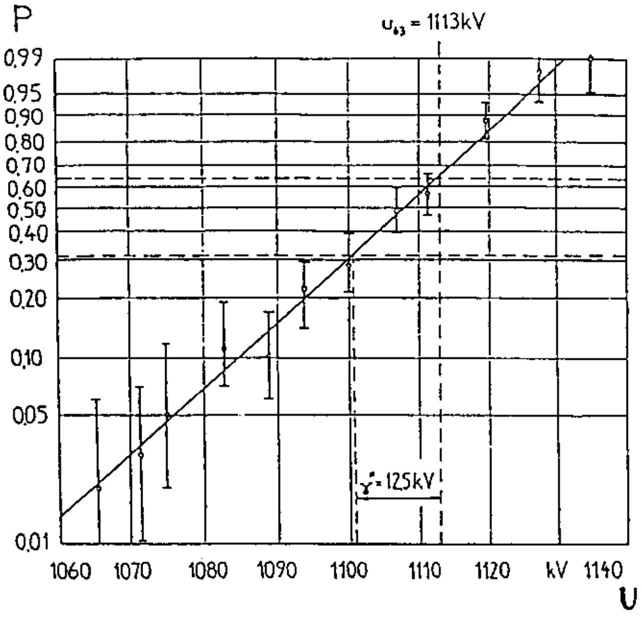


Fig. 2 (A8) - Graphical analysis of Class-1 test (MLM). Gumbel paper. Empirical discharge probability function P(U).

to approximate the empirical distribution function. A conventional linear regression (e.g. minimum squares) is applied to the results of this transformation. The parameters of the empirical function are calculated from the regression coefficients. For the evaluation of the confidence limits refer to the literature.

The goodness-of-fit test. The test consists of checking of how many of the n confidence intervals $P_{li} \longmapsto P_{ui}$ of the measured frequencies f_i are crossed by the straight line representing the approximating function. If this line does not cross at least C*100% of them, the test is considered not passed, and the approximation should be rejected.

Example. A class-1 test (MLM) was performed with n = 12, and m_j = 100 for all levels (j = 1, 2, .., 20). The results are given in Table 2 (A8). The independence for each level was tested, as previously described, using the actual sequence of discharges and withstands. A Gumbel distribution was adopted as reaction function. The measured frequencies and their 0.95 confidence intervals were plotted at the relevant voltage levels U_j on Gumbel paper (Fig. 2 (A8)). The best approximation (i.e. the straight line in Fig. 2 (A8) was calculated by linear regression of the transformed frequencies. The goodness-of-fit test is satisfied, because the straight line crosses at least 19 confidence intervals (C*n = 0.95*20 = 19. It actually crosses all the n = 20 intervals).

Analysis of Class-2 tests (UDM)

The independence test. The trend of the outcomes U_j (j= 1, 2, ..., n) is compared with their arithmetic mean value $\hat{U}$. If they scatter randomly around $\hat{U}$, the independence cannot be rejected. An increasing or decreasing tendency characterizes a dependency, and the test should be repeated under improved conditions.

Table 2 (A8) - Result of a Class-1 test.

level j	*voltage* U_j (kV)	*impulses* m_j	*discharges* k_j	*frequency* f_j
1	1065	100	2	0.02
2	1071	100	3	0.03
3	1075	100	5	0.05
4	1083	100	11	0.11
5	1089	100	10	0.10
6	1094	100	21	0.21
7	1100	100	29	0.29
8	1107	100	48	0.48
9	1111	100	56	0.56
10	1120	100	88	0.88
11	1128	100	98	0.98
12	1135	100	99	0.99

No graphical representation is foreseen, as no approximating function is necessary for the analysis. However, the sequence of test results shown in Fig. 1 (A8) can be considered a kind of graphical representation.

The analytical evaluation. It is performed according to IEC Publication 60 [1] [2]. At the end of the test the n voltages U_j result distributed in m voltage levels Ui (i = 1, ,,, m << n), at each of which k_i series of test voltages were applied. The expected value U_p can be simply estimated by the arithmetic mean value of the levels U_i (weighted with k_i) at which at least two series of test voltages have been applied ($k_j \geq 2$):

$$U_p = \frac{\sum_{i=1}^{m} (k_i * U_i)}{\sum_{i=1}^{m} (k_i)} \qquad \text{(A8.3)}$$

More precise estimations of the value U_p, of its standard deviation σ, and of the relevant confidence limits are given in the literature.

No goodness-of-fit test is required, for the same reason put forward for the graphical representation.

It is worth mentioning that with the classical methods all estimation of UDM are related to the Gauss distribution. Furthermore, with the classical methods, it is essential that no mistake is made during the test (increasing the voltage instead decreasing it, applying the test voltage at a wrong level, etc.).

Example. Fig. 3-A (A8) reports the results of a class-2 test (UDM) with procedure A, for h = 7 impulses per level. Consequently,

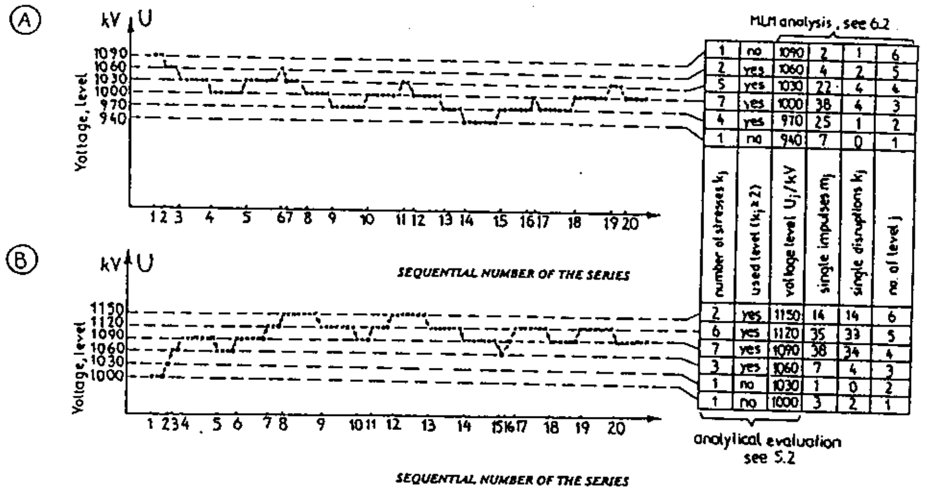


Fig. 3 (A8) - Principle of Class-2 test (UDM) and relevant evaluation.

- *Procedure A (for expected p'=10% ≤ 50%) with h = 7 impulses per level.*
- *Procedure B (for expected p'=90% ≥ 50%) with h = 7 impulses per level.*

p' = 0.10. From equation (A8.3) U_{10} = 1008.3 kV is obtained. Fig. 3-B (A8) shows a test with procedure B: p' = 0.90, U_{90} = 1101.7 kV.

Analysis of Class-3 tests (PSM)

The independence test. A simple trend test can be performed as described in the analysis of Class-2 tests (Fig. 4 (A8)). An improvement of this test is the subdivision of the samples into groups and the comparison of the parameters of the first and the last group with each other and with the whole sample by statistical tests. An additional independence test consists of the reduction to a series of alternating events such as "higher" and "lower" than the mean value, and the application to that of the "runs test".

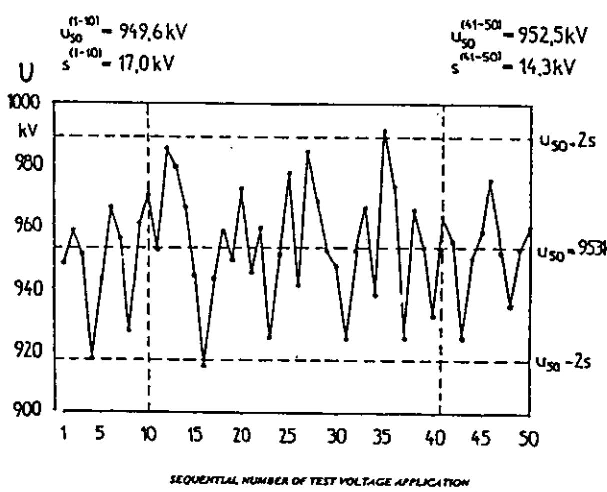


Fig. 4 (A8) - Sequential results of Class-3 test. Relevant independence test.

The graphical representation. All n outcomes should be put in order of increasing values ($x_1 < x_2 < .. < x_n$) and for each value x_i the cumulative frequency $F_i = F(x_i) = i/n$ is determined, and used for the construction of a "stair" function on the relevant probability paper (Fig. 5 (A8)). Each outcome causes a step of height $p_s = 1/n$. This representation can be approximated by a polygon connecting the points of coordinates x_i, $i/(n+1)$ or x_i, $(i-0.5)/n$. The best fitting straight line may be drawn by eye. However, to check if the approximation is sufficient, it is suggested to use the analytical expression of the best fitting function later described.

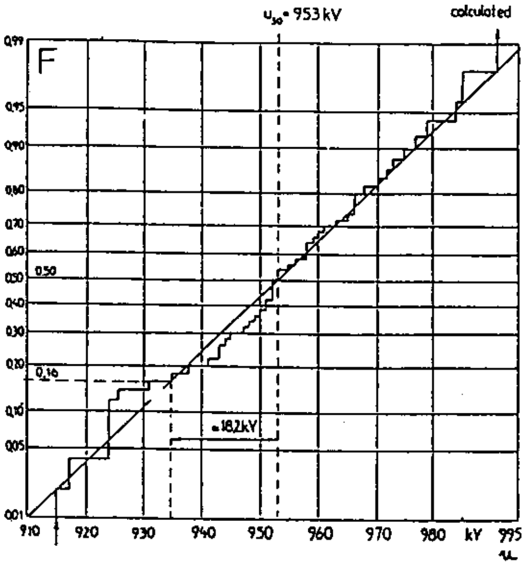


Fig. 5 (A8) - Graphical analysis of Class-3 test (PSM). Gauss paper. Empirical cumulative function F(u).

ytical evaluation. It is based on estimators: the mean and the meaiation (x_m,s_x) of the measured valumbel and Gauss, or (y_m,s_y) of their s, for Weibull functions.

metic mean and measured deviation lues x_i are:

$$\sum_{i=1}^{n}(x_i)\,; \qquad \overline{\frac{\sum_{i=1}^{n}(x_i-x_m)^2}{(n-1)}} \qquad (A8.4)$$

the Gauss distribution the point s are:

$$_n; \qquad z^* = s_x \qquad (A8.5)$$

lation of the confidence limits is led, and for the Gumbel distribution estimators are:

$$\frac{x}{1}; \qquad \eta^* = x_m+\gamma^* k_{n2} \qquad (A8.5)$$

and k_{n2} are correction factors given 3 (A8).

ithmic mean value and measured def the values x_i are:

$$\sum_{i=1}^{n}(\ln(x_i))\,; \qquad \overline{\frac{\sum_{i=1}^{n}(\ln(x_i)-y_m)^2}{(n-1)}} \qquad (A8.7)$$

the Weibull distribution (two pax_o = 0) the point estimators are:

$$\delta^* = 1.283\,\frac{k_{n3}}{s_y}\,; \qquad \eta^* = e^{(y_m+0.45 s_y)} \qquad (A8.8)$$

where k_{n3} is a correction factor given in Table 3 (A8).

Table 3 (A8) - Correction factors for the point estimators of the Gumbel and Weibull distributions.

Sample size	Distributions		
	Gumbel		Weibull
n	k_{n1}	k_{n2}	k_{n3}
10	0.950	0.495	0.863
15	1.021	0.513	0.906
20	1.063	0.524	0.923
30	1.112	0.536	0.950
40	1.141	0.544	0.961
50	1.161	0.549	0.969
60	1.175	0.552	0.974
80	1.194	0.557	0.980
100	1.207	0.560	0.984
limit	1.282	0.577	1.000

For the evaluation of the Weibull distribution with three parameters ($x_o \neq 0$), and for the calculation of the confidence intervals, reference is made to the literature or, still better, the use of the likelihood procedure is recommended.

The goodness-of-fit test. It can be based on the difference between the calculated analytical approximation and the empirical function previously described.

Example. The result of a class-3 test (PSM) is reported in Fig. 4 (A8), the independence was not rejected by the two comparisons: of the single values with the arithmetic mean value U_{50}, and of the first and last ten values. As frequency function, a Gauss distribution was assumed. The "stair" function was plotted on Gauss paper (Fig. 5 (A8)). The parameters U_{50} and z were determined according to equation (A8.5), and the approximation function was plotted in Fig. 5 (A8) as a straight line. The approximation is not rejected, because the straight line fits the stair function very well.

Le CIGRÉ a apporté le plus grand soin à la réalisation de cette brochure thématique numérique afin de vous fournir une information complète et fiable.

Cependant, le CIGRÉ ne pourra en aucun cas être tenu responsable des préjudices ou dommages de quelque nature que ce soit pouvant résulter d'une mauvaise utilisation des informations contenues dans cette brochure.

Publié par le CIGRÉ
21, rue d'Artois
FR-75 008 PARIS
Tél. : +33 1 53 89 12 90
Fax : +33 1 53 89 12 99

Copyright © 2000

Tous droits de diffusion, de traduction et de reproduction réservés pour tous pays.

Toute reproduction, même partielle, par quelque procédé que ce soit, est interdite sans autorisation préalable. Cette interdiction ne peut s'appliquer à l'utilisateur personne physique ayant acheté ce document pour l'impression dudit document à des fins strictement personnelles.
Pour toute utilisation collective, prière de nous contacter à sales-meetings@cigre.org

The greatest care has been taken by CIGRE to produce this digital technical brochure so as to provide you with full and reliable information.

However, CIGRE could in any case be held responsible for any damage resulting from any misuse of the information contained therein.

Published by CIGRE
21, rue d'Artois
FR-75 008 PARIS
Tel : +33 1 53 89 12 90
Fax : +33 1 53 89 12 99

Copyright © 2000

All rights of circulation, translation and reproduction reserved for all countries.

No part of this publication may be produced or transmitted, in any form or by any means, without prior permission of the publisher. This measure will not apply in the case of printing off of this document by any individual having purchased it for personal purposes.
For any collective use, please contact us at sales-meetings@cigre.org