

TECHNICAL REPORT

**Guide for the statistical analysis of ageing test data –
Part 2: Validation of procedures for statistical analysis of censored normally
distributed data**



THIS PUBLICATION IS COPYRIGHT PROTECTED

Copyright © 2010 IEC, Geneva, Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either IEC or IEC's member National Committee in the country of the requester.

If you have any questions about IEC copyright or have an enquiry about obtaining additional rights to this publication, please contact the address below or your local IEC member National Committee for further information.

Droits de reproduction réservés. Sauf indication contraire, aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie et les microfilms, sans l'accord écrit de la CEI ou du Comité national de la CEI du pays du demandeur.

Si vous avez des questions sur le copyright de la CEI ou si vous désirez obtenir des droits supplémentaires sur cette publication, utilisez les coordonnées ci-après ou contactez le Comité national de la CEI de votre pays de résidence.

IEC Central Office
3, rue de Varembe
CH-1211 Geneva 20
Switzerland
Email: inmail@iec.ch
Web: www.iec.ch

About IEC publications

The technical content of IEC publications is kept under constant review by the IEC. Please make sure that you have the latest edition, a corrigenda or an amendment might have been published.

- Catalogue of IEC publications: www.iec.ch/searchpub

The IEC on-line Catalogue enables you to search by a variety of criteria (reference number, text, technical committee,...). It also gives information on projects, withdrawn and replaced publications.

- IEC Just Published: www.iec.ch/online_news/justpub

Stay up to date on all new IEC publications. Just Published details twice a month all new publications released. Available on-line and also by email.

- Electropedia: www.electropedia.org

The world's leading online dictionary of electronic and electrical terms containing more than 20 000 terms and definitions in English and French, with equivalent terms in additional languages. Also known as the International Electrotechnical Vocabulary online.

- Customer Service Centre: www.iec.ch/webstore/custserv

If you wish to give us your feedback on this publication or need further assistance, please visit the Customer Service Centre FAQ or contact us:

Email: csc@iec.ch
Tel.: +41 22 919 02 11
Fax: +41 22 919 03 00

TECHNICAL REPORT

**Guide for the statistical analysis of ageing test data –
Part 2: Validation of procedures for statistical analysis of censored normally
distributed data**

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

PRICE CODE **XG**

ICS 29.035.01

ISBN 2-8318-1086-2

CONTENTS

FOREWORD.....	4
INTRODUCTION.....	6
1 Scope.....	7
2 Normative references	7
3 Terms, symbols and definitions	7
3.1 Definitions	7
3.2 Symbols	10
4 Statistical calculations	11
4.1 General.....	11
4.1.1 Objectives of the analysis.....	11
4.1.2 Outline of work program	11
4.1.3 Computer programs.....	12
4.2 Calculations for complete (uncensored) data	12
4.2.1 Sub-group	12
4.2.2 Group (several sub-groups).....	12
4.2.3 The F-function	13
4.2.4 Bartlett's χ^2 function	13
4.2.5 Student's t function.....	13
4.2.6 Student's t function for the difference of two means.....	14
4.3 Censored data groups	14
4.3.1 Sub-groups.....	15
4.3.2 Group (several sub-groups) – Analysis of variance	15
4.3.3 Group (several sub-groups) – Analysis of covariance (regression analysis).....	16
4.3.4 Difference of means of 2 sub-groups	18
4.4 Complete (uncensored) data	19
5 Random number generation	19
5.1 General.....	19
5.2 Uniform variates	19
5.3 Normal variates	20
6 Validation procedures.....	20
6.1 Size of samples.....	20
6.2 Validation of random number generation	21
7 Outline of tasks	21
7.1 Optimum size of samples	21
7.2 Uniformly and normally distributed random number generation.....	22
7.3 Examine properties of single sub-groups	22
7.3.1 Correlation of mean and standard deviation (SD).....	22
7.3.2 Distribution and possible bias of mean and SD estimates	22
7.4 Examine properties of compound groups	22
7.5 Procedure for testing difference of two means	22
8 Discussion of results and conclusions	22
8.1 Optimum size of samples	22
8.2 Correlation coefficients of mean and standard deviation of sub-groups.....	22
8.3 Values of mean, variance and SD of mean of censored data sub-groups.....	22
8.4 Distribution functions of mean and variance of censored data sub-groups.....	23

8.5	Properties of compound data groups	23
8.5.1	F (variance ratio) function	23
8.5.2	“Student’s” t function	23
8.5.3	Bartlett’s χ^2 (variance equality) function	24
8.6	Generation of uniformly distributed random numbers	24
8.7	Values of the t -ratio for the difference of two means	24
9	Conclusions	25
9.1	Random number generation	25
9.1.1	Uniform distribution	25
9.1.2	Normal distribution	25
9.2	Statistical functions of censored data groups	25
9.2.1	Simple sub-groups	25
9.2.2	Compound groups – Analysis of variance	25
9.2.3	Compound groups – Regression analysis	26
9.2.4	Difference of means of 2 sub-groups	26
10	Summary of conclusions and recommendations	27
10.1	Summary of conclusions	27
10.2	Recommendations	27
	Annex A (informative) Results	28
	Annex B (informative) Key to graph features	43
	Annex C (informative) Graphical displays for the group statistical functions	47
	Annex D (informative) Computer programs	148
	Annex E (informative) Worked example	150
	Bibliography	151
	Figure A.1 – Confidence limits for value counts	29
	Table A.1 – Confidence limits for value counts	28

INTERNATIONAL ELECTROTECHNICAL COMMISSION

**GUIDE FOR THE STATISTICAL ANALYSIS
OF AGEING TEST DATA –****Part 2: Validation of procedures for statistical analysis
of censored normally distributed data****FOREWORD**

- 1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
- 2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
- 3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
- 4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
- 5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.
- 6) All users should ensure that they have the latest edition of this publication.
- 7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
- 8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
- 9) Attention is drawn to the possibility that some of the elements of this IEC Publication may be the subject of patent rights. IEC shall not be held responsible for identifying any or all such patent rights.

The main task of IEC technical committees is to prepare International Standards. However, a technical committee may propose the publication of a technical report when it has collected data of a different kind from that which is normally published as an International Standard, for example "state of the art".

IEC 60493-2 which is a technical report, has been prepared by IEC technical committee 112: Evaluation and qualification of electrical insulating materials and systems.

The text of this technical report is based on the following documents:

Enquiry draft	Report on voting
112/140/DTR	112/145/RVC

Full information on the voting for the approval of this technical report can be found in the report on voting indicated in the above table.

This publication has been drafted in accordance with the ISO/IEC Directives, Part 2.

This publication contains attached files in the form of a CD-ROM. These files are intended to be used as a complement and do not form an integral part of the standard.

A list of all parts of the IEC 60493 series, published under the general title *Guide for the statistical analysis of ageing test data*, can be found on the IEC website.

The committee has decided that the contents of this publication will remain unchanged until the stability date indicated on the IEC web site under "<http://webstore.iec.ch>" in the data related to the specific publication. At this date, the publication will be

- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

A bilingual version of this publication may be issued at a later date.

INTRODUCTION

Procedures for estimating ageing properties are described in specific test procedures, or are covered by the general documents on test procedures for ageing tests with a specific environmental stress (e.g. temperature, radiation, partial discharges).

In many cases, a certain property is determined as a function of time at different ageing stresses, and a time to failure based on a chosen end-point criterion is found at each ageing stress. A plot of time to failure versus ageing stress may be used to obtain an estimate of the time to failure for similar specimens exposed to a specified stress, or to obtain an estimate of the value of stress which will cause failure in a specified time.

The physical and chemical laws governing the ageing phenomena may often lead to the assumption that a linear relationship exists between the property examined and the ageing time at fixed ageing stresses, or between certain mathematical functions of property and ageing time, e.g. square root or logarithm.

The relationships between property values, ageing time, ageing and ageing exposure temperature are determined by mathematical procedures known as analysis of variance, analysis of covariance and regression analysis. In some cases, the property values cannot be determined because of time or property measurement limitations. Such data are referred to as “censored data”.

The mathematical procedures are well known and accepted as valid for complete (uncensored) data. This technical report describes the work done to validate the procedures for censored data.

GUIDE FOR THE STATISTICAL ANALYSIS OF AGEING TEST DATA –

Part 2: Validation of procedures for statistical analysis of censored normally distributed data

1 Scope

This part of IEC 60493 provides an account of the work done in designing and validating the statistical procedures for operations on censored groups of normally distributed data.

The relationship to similar operations on complete (uncensored) data is examined, and it is shown that “mixed” or wholly uncensored data groups may be analysed in the same way as wholly censored ones.

Attention is drawn to the effect of some variation of group size and extent of censoring, as well as the effect of non-uniform data group sizes.

2 Normative references

IEC 60216-3, *Electrical insulating materials – Thermal endurance properties – Part 3: Instructions for calculating thermal endurance characteristics*

IEC 60216-5, *Electrical insulating materials – Thermal endurance properties – Part 5: Determination of relative thermal endurance index (RTE) of an insulating material*

IEC 60216-6, *Electrical insulating materials – Thermal endurance properties – Part 6: Determination of thermal endurance indices (TI and RTE) of an insulating material using the fixed time frame method*

IEC 60493-1, *Guide for the statistical analysis of ageing test data. Part 1: Methods based on mean values of normally distributed test results*

3 Terms, definitions and symbols

3.1 Definitions

For the purposes of this document, the following terms and definitions apply.

3.1.1

ordered data

a set of data arranged in sequence so that in the appropriate direction through the sequence each member is greater than or equal to its predecessor.

NOTE “Ascending order” in this report implies that the data is ordered in this way, the first being the smallest.

3.1.2

order-statistic

each individual value in a set of ordered data identified by its numerical position in the sequence

3.1.3

incomplete data

ordered data, where the values above and/or below defined points are not known

3.1.4

censored data

incomplete data, where the number of unknown values is known

NOTE If the censoring is begun above/below a specified numerical value, the censoring is Type I. If above/below a specified order-statistic it is Type II. This technical report is concerned only with Type II.

3.1.5

truncated data

incomplete data where the number of unknown values is not known

NOTE This report is not concerned with truncated data.

3.1.6

degrees of freedom

DF

the number of data values minus the number of parameter values

3.1.7

variance of a data group

the sum of the squares of the deviations of the data from a reference level defined by one or more parameters, for example a mean value (1 parameter) or a line (2 parameters, slope and intercept), divided by the number of degrees of freedom

3.1.8

standard deviation

SD

the square root of the variance

3.1.9

central second moment of a data group

the sum of the squares of the differences between the data values and the value of the group mean divided by the number of data in the group

3.1.10

covariance of data groups

for two groups of data with equal numbers of elements where each element in one group corresponds to one in the other, the sum of the products of the deviations of the corresponding members from their group means, divided by the number of degrees of freedom

3.1.11

regression analysis

the process of deducing the best-fit line expressing the relation of corresponding members of two data groups by minimizing the sum of squares of deviations of members of one of the groups from the line

NOTE The parameters are referred to as the regression coefficients.

3.1.12

correlation coefficient

a number expressing the completeness of the relation between members of two data groups. In value it is equal to the covariance divided by the square root of the product of the variances of the groups. The value of its square is between 0 (no correlation) and 1 (complete correlation)

3.1.13**primary statistical function**

a statistical function of a data sub-group (for example, mean or variance), obtained using the Saw coefficients

3.1.14**Saw coefficient**

one of the coefficients developed by J.G.Saw for calculating the primary statistical functions of a single sub-group

NOTE There are four (4) coefficients used in this report. Saw originally defined five (5), the fifth being intended for estimating the variance of the variance estimate.

3.1.15**secondary statistical function**

a statistical function of a data group, calculated from the primary functions of the sub-groups

NOTE For example, the Fisher variance ratio function, the t function or the χ^2 function.

3.1.16**data sub-group**

a single set of data which may be used with other sub-groups to form a compound group

3.1.17**data group**

a set of data comprising more than one sub-group, of which the secondary statistical functions are to be calculated

3.1.18**group sample**

the set of data groups assembled to establish the distribution of one or more secondary statistical functions.

NOTE The size of the sample is normally at least 10^6

3.1.19**parent population**

randomly selected data for the groups in a sample from an effectively infinite population of normally distributed numbers.

NOTE The size of this population is such that removal of the sample data does not significantly affect the size of the remaining population

3.1.20**analysis of variance****ANOVA**

separation of the components of the variance of a data group, so that further calculations, (e.g. linear regression) may be executed

3.1.21**RAN1**

algorithm for generation of uniformly distributed random numbers using the linear congruential method

3.1.22**RAN3**

algorithm for generation of uniformly distributed random numbers using the subtractive method

3.1.23

unbiased estimate (of a statistical property)

the estimate of a property of a censored data group, such that the mean value of a very large number of estimates is equal to the mean value of the property of groups drawn from the parent normally distributed population

3.2 Symbols

In this technical report the following symbols are used:

Symbol	Definition	Equation at the first citation	Clause
$x_{i,j}$	Value numbered j in sub-group i	1	4.2.1
i, j	Indices for x	1	4.2.1
$\bar{x}_i$	x -mean in sub-group i	1	4.2.1
n_i	number of values in sub-group i	1	4.2.1
σ_i^2	variance of sub-group i	2	4.2.1
$s_{\bar{x}_i}^2$	variance of x -mean in sub-group i	3	4.2.1
N	total number of values in Group	4	4.2.2
k	number of sub-groups	4	4.2.2
$\bar{\bar{x}}$	general mean of group	5	4.2.2
σ_N^2	numerator variance	6	4.2.3
σ_D^2	denominator variance	7	4.2.3
F	variance ratio	8	4.2.3
σ_T^2	overall variance: $(N-1)$ DF	9	4.2.3
c	Intermediate constant for χ^2	10	4.2.4
χ^2	value of χ^2 for equality of variances test	11	4.2.4
t	"Student's" t ratio	12	4.2.5
$\bar{\bar{x}}_c$	confidence limit of group general mean	13	4.2.5
μ	"Saw" coefficient for sub-group mean	16	4.3.1
α	"Saw" coefficient for sub-group variance	17	4.3.1
β	"Saw" coefficient for sub-group variance	17	4.3.1
ε_i	"Saw" coefficient for variance of sub-group mean	18	4.3.1
a	adjustment to t or χ^2	39	4.3.3
π	constant 3.14159.....	50	5.3
ν_1	temporary variable	54	

Symbol	Definition	Equation at the first citation	Clause
v_2	temporary variable	54	5.3
P	cumulative binomial probability	55	6.1
p	probability per trial	55	6.1

4 Statistical calculations

4.1 General

4.1.1 Objectives of the analysis

In this section, the statistical calculations are outlined. The calculations are basically those of conventional “analysis of variance” (ANOVA), and cover the derivation of the functions necessary for that analysis. The ANOVA calculations are also relevant with slight modification to regression analysis as required for IEC 60216-3 and IEC 60493-1.

The analysis is carried out on a “group” of data, made up of k ($k \geq 1$) “sub-groups”. Each of these comprises several normal variate values. In general, they are not equal in size. All or none may be censored. The data for the groups are derived from a population which is nominally “normalised Gaussian”, i.e. having mean equal to 0 and variance 1.

Calculation equations are given for the case where none of the data is censored (i.e. all sub-groups are “complete”) and for the other case, where one or more of the sub-groups is censored. The extent of censoring need not be the same for all sub-groups.

4.1.2 Outline of work program

The first decision was that an analytical determination of the statistical distributions of the various functions was beyond the capacity of the author of this report, and that “Monte Carlo” techniques should be employed. (This means generating numbers of groups of normally distributed pseudo-random numbers, calculating the required functions of those groups and comparing their distributions with the theoretical distributions of the functions e.g. F and t).

In view of the very large number of test groups of any one kind which would be required to ensure adequate confidence in the result (this has been estimated at about 1 million), it was decided that the first step must be validation of the random number generator in the computer used for the whole exercise.

Generation of normally distributed random numbers was achieved by first generating uniformly distributed numbers, and converting pairs of these to give the normally distributed ones. The validation was then executed by generating the required functions from complete groups of random numbers, and comparing their distributions with the theoretical.

When this had been done satisfactorily (some procedures commonly used were quite unsatisfactory for the very large populations of random numbers needed) it was possible to proceed with generating the functions of censored data groups.

NOTE In one case, the number of uniformly distributed random numbers required exceeded 2×10^9 .

The first stage was validation of the primary statistical functions directly generated from the “Saw” coefficients. These are the estimates of mean, variance of data and variance of mean estimate of data sub-groups.

After acceptance of this validity, the distributions of functions of groups of multiple sub-groups were generated for comparison with theory:

- variance ratio for analysis of variance (ANOVA) F ;
- student's t for ANOVA;
- variance equality (χ^2) for ANOVA;
- variance ratio for linear regression F ;
- student's t for linear regression;
- variance equality (χ^2) for linear regression;
- regression slope: mean and variance;
- regression intercept: mean and variance;
- student's t for comparison of the estimates of the means of two sub-groups;

4.1.3 Computer programs

The programs employed in the ANOVA and linear regression sections are provided on a CDROM with this report. Both executable versions and source code are included. This enables the user of the report to repeat some of the calculations, and also to carry out ANOVA or linear regression calculations on single data groups. In addition, the routines used for generation of normal and uniform random numbers may be found useful. Further details are included in a file on the disc.

4.2 Calculations for complete (uncensored) data

The calculations for complete data groups are primarily intended for the appraisal of the data generation methods.

4.2.1 Sub-group

$$\text{Subgroup mean} \quad \bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} \quad (1)$$

$$\text{Subgroup variance} \quad \sigma_i^2 = \frac{\left[\sum_{j=1}^{n_i} x_{ij}^2 - n_i \bar{x}_i^2 \right]}{(n_i - 1)} \quad (2)$$

$$\text{Variance of mean of sub-group} \quad s_{\bar{x}_i}^2 = \frac{\sigma_i^2}{n} \quad (3)$$

4.2.2 Group (several sub-groups)

It should not be necessary to emphasise that for the purposes of this analysis, all groups in any one sample shall have the same structure. The sub-group sizes and their degree of censoring shall not vary between groups but all sub-groups within a group need not have the same size or degree of censoring.

The primary statistical functions are the means and variances of the sub-groups. From these are derived the three (3) functions required for the analysis of variance. These are:

- variance ratio: F ("Fisher" test), also used as a linearity test;
- "Bartlett's" χ^2 test: for equality of variances;
- "student's" t : used for calculating confidence intervals of the general mean estimate.

For the calculation of these properties, the total number of values in a group and the general mean of all values in the group are required.

$$\text{Total number of values in Group} \quad N = \sum_{i=1}^k n_i \quad (4)$$

$$\text{General mean of all values in Group} \quad \bar{\bar{x}} = \frac{\sum_{i=1}^k \bar{x}_i n_i}{N} \quad (5)$$

4.2.3 The F-function

The F-test is based on a ratio and is used to assess whether the variance of the sub-group means (numerator) is greater than would be expected, based on the merged values of the sub-group variances (denominator). The numerator and denominator variances are calculated by Equations 6 and 7. The statistical distribution of the ratio is well known, and may be obtained from tabulated values or by the use of a standard algorithm.

$$\text{Numerator variance for F:(k-1) DF} \quad \sigma_N^2 = \frac{\left[\sum_{i=1}^k n_i \bar{x}_i^2 - N \bar{\bar{x}}^2 \right]}{(k-1)} \quad (6)$$

$$\text{Denominator variance for F:(N-k) DF} \quad \sigma_D^2 = \frac{\left[\sum_{i=1}^k \sigma_i^2 (n_i - 1) \right]}{(N-k)} \quad (7)$$

$$\text{Variance ratio} \quad F = \frac{\sigma_N^2}{\sigma_D^2} \quad (8)$$

The overall variance is a combination of the above (Equations 6 and 7), and is used to estimate the confidence intervals of a general mean estimate using the *t*-function of Equation 12

$$\text{Overall variance:(N-1) DF} \quad \sigma_T^2 = \frac{\left[\sum_{i=1}^k n_i \bar{x}_i^2 + \sum_{i=1}^k \sigma_i^2 (n_i - 1) - N \bar{\bar{x}}^2 \right]}{(N-1)} \quad (9)$$

4.2.4 Bartlett's χ^2 function

The Bartlett's χ^2 function provides a test of whether the variances of the sub-groups can be considered not to be significantly different

$$\text{Intermediate constant for } \chi^2 \quad c = 1 + \frac{\left[\sum_{i=1}^k \frac{1}{(n_i - 1)} - \frac{1}{(N - k)} \right]}{3(k-1)} \quad (10)$$

$$\text{Variance equality} \quad \chi^2 = \frac{1}{c} \left[(N-k) \log(\sigma_D^2) - \sum_{i=1}^k (n_i - 1) \log(\sigma_i^2) \right] \quad (11)$$

4.2.5 Student's *t* function

$$\text{"Student's" } t \text{ function} \quad t = \bar{\bar{x}} \sqrt{\frac{N}{\sigma_T^2}} \quad (12)$$

$$\text{Confidence limit of } \bar{\bar{x}} \quad \bar{\bar{x}}_C = \bar{\bar{x}} + t_{p, n-1} \quad (13)$$

For Equation 13, *p* is the probability to be applied to the *t*-function with *N*-1 DF.

4.2.6 Student's t function for the difference of two means

The Student's t function is used to assess the significance of a difference between the observed and expected values of the general mean, and also to give values for the confidence levels of the general mean.

For the special case of two sub-groups, the value of t (not the same as in Equation 12) for the significance of the difference between the means is calculated using the equation:

$$t = \pm \sqrt{F} \quad (14)$$

Here F has been calculated by Equation 8. The value of t is given the same sign as $(\bar{x}_1 - \bar{x}_2)$ and has $(N - 2)$ degrees of freedom.

An identical result may be obtained from the conventional Equation 15 for this difference, but Equation (14) is more convenient if F has already been calculated:

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{((n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2)}{(n_1 + n_2 - 2)} \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (15)$$

4.3 Censored data groups

Much of the work on statistical treatment of normally distributed censored data has been done using "maximum likelihood" methods (see, for example, Reference [6]²⁾). These generally involve serious computer activity, and were not thought to be suitable for inclusion in a standard such as 216. The use of distributions such as Weibull and Gumbell is also unsuitable, since there appears to be no way in which derived statistical functions (F , χ^2 and t) for groups comprising several sub-groups can be calculated.

In about 1960, a different method was developed by J.G.Saw, which involved simple algebraic procedures, but which required difficult and tedious calculations to obtain the necessary coefficients for normally distributed data (Reference [1]). This work was largely ignored.

The situation changed with the advent of more powerful personal computers, which enabled the calculation of the coefficients needed and these are tabulated in IEC 60216-3. At the same time, it became feasible to use Monte Carlo methods to obtain values for the statistical distributions of the desired statistical functions. With the personal computers available today, it is feasible to conduct Monte Carlo simulations with quite large data groups (for example 4 sub-groups of 200 numbers) with a sample of 10^6 or even 10^7 groups.

The calculation procedures developed by Saw enabled simple calculation of three primary functions, mean, variance and variance of mean. Although these results were unbiased estimators of the functions of the notional parent population of the data, there was no way of establishing their actual statistical distributions. In addition, these functions are of little use in isolation: for example, a value of the mean is only of use if its relation to a desired value can be established.

The statistical tests in IEC 60216-3, IEC 60216-5, and IEC 60216-6 are required to establish (1) whether all groups of data can be considered to belong to the same parent statistical distribution (2) whether the scatter of the mean values themselves or their deviations from a desired pattern (e.g. regression line) is compatible with the scatter of the individual data and (3) whether the uncertainty in the general mean of the data is acceptably small. These tests involve secondary statistical functions derived from the mean, variance and variance of the mean. These functions, and the tests based on them are, respectively:

2) Figures in square brackets refer to the Bibliography

- Bartlett's χ^2 (Variance equality test);
- F (Fisher or Variance ratio test);
- t (Student's t -test).

Similar functions can be derived from the Saw primary functions. Although the statistical distributions of these functions is not known, it is possible to obtain this information empirically by Monte Carlo methods if a sufficiently accurate source of normally distributed random numbers is available, and the establishment of a suitable computer program for generating “pseudo-random” normally distributed data is a primary task.

It must be emphasised that the calculations using these functions are based on the assumption that the initial data are distributed in the normal or Gaussian manner. However, since the three tests are widely used uncritically without testing for normality of complete or uncensored data, and usually without serious error it may be expected that the corresponding tests on censored data will be similarly useful.

The values of the coefficients μ , α , β and ε are tabulated in IEC 60216-3 for group sizes up to 31, censored at any point above the median value. As a separate exercise, values for group sizes up to 200 are available.

4.3.1 Sub-groups

$$\text{Mean of sub-group} \quad \bar{x} = (1 - \mu)x_n + \frac{\mu}{(n-1)} \sum_{j=1}^{n-1} x_j \quad (16)$$

$$\text{Variance of sub-group} \quad s^2 = \alpha \sum_{j=1}^{n-1} (x_n - x_j)^2 + \beta \left[\sum_{j=1}^{n-1} (x_n - x_j) \right]^2 \quad (17)$$

$$\text{Variance of mean of sub-group} \quad s_{\bar{x}}^2 = \frac{\varepsilon s^2}{n} \quad (18)$$

NOTE 1 In the original paper by Saw [1], the above definition of ε was not used. The multiplier value for the variance of mean of a sub-group was ε/n .

4.3.2 Group (several sub-groups) – Analysis of variance

Each group generated comprises a number of sub-groups, which may be complete or censored. The size of sub-group and fraction of data censored need not be the same for all sub-groups in the same group.

$$\text{Factor for variance of general mean} \quad \varepsilon = \frac{\sum_{i=1}^k \varepsilon_i}{k} \quad (19)$$

$$\text{Numerator variance for F} \quad \sigma_N^2 = \frac{\left[\sum_{i=1}^k n_i \bar{x}_i^2 - N \bar{\bar{x}}^2 \right]}{(k-1)} \quad (20)$$

$$\text{Denominator variance for F} \quad \sigma_D^2 = \frac{\varepsilon \left[\sum_{i=1}^k \sigma_i^2 (n_i - 1) \right]}{(N - k)} \quad (21)$$

$$\text{Overall variance} \quad \sigma_T^2 = \frac{\left[\sum_{i=1}^k n_i \bar{x}_i^2 + \varepsilon \sum_{i=1}^k \sigma_i^2 (n_i - 1) - N \bar{\bar{x}}^2 \right]}{(N - 1)} \quad (22)$$

The derived functions F , χ^2 and t are calculated as in 4.2.2, using the above values of the primary functions.

4.3.3 Group (several sub-groups) – Analysis of covariance (regression analysis)

The procedures of IEC 60216-3 are those of analysis of variance, covariance and regression, which will be briefly explained as follows:

The work on statistical analysis of censored data so far reported is an extension of the classical one-way analysis of variance procedure.

If the data of the sub-groups are represented by x , each sub-group is associated with another value, z .

This association is most simply represented in vector notation:

$$x = a + bz + sr \quad \text{Each individual datum value is} \quad x_{ij} = a + bz_i + sr_j$$

Bold typeface (e.g. z) implies an array (vector) of values, normal typeface (e.g. a) implies a single valued (scalar) number.

x is the array of data under examination (log t_{ij} in IEC 60216)

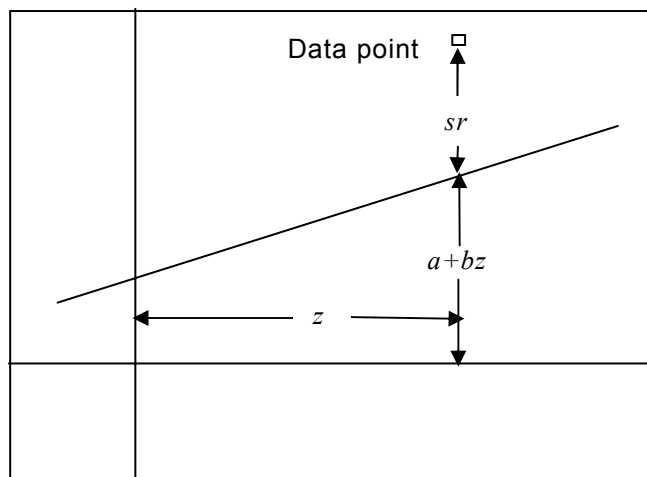
z is the array of values of an “independent” variable ($1/(\vartheta + \vartheta_0)$ in IEC 60216)

r is the array of normalized Gaussian random numbers (mean = 0, SD = 1)

a and b are the intercept and slope parameters of a linear equation

s is the SD of the parent population of x

This is illustrated by the drawing below.



The above differs from that of (4.3.2) in that each subgroup is associated with a value of another variable, referred to in this section as z_i . The objective of the analysis is to determine whether there is a linear relationship between x and z and, if so, its parameters and properties.

The parameters and properties in question are:

- slope (b) and intercept (a) of regression line;
- equality of variance of subgroups (χ^2);
- linearity of regression (F).

The following is an outline of the calculation sequence.

Where appropriate, the symbols of Section 4.3.1 etc are used.

For these calculations:

Values of y_{ij} are for the actual values of variable j in subgroup i

n_i = number of values known in subgroup i

m_i = total number in subgroup i

α_i, β_i, μ_i and ε_i are the “Saw” coefficients for these values of m and n

k = number of subgroups

c = intermediate value for χ^2 calculation

A = adjustment factor for χ^2 calculation

b and a are the slope and intercept of the regression line

$t_{p,n-1}$ = tabulated value of t for probability p and $n-1$ degrees of freedom

$$\text{Calculate the total number of values} \quad M = \sum_{i=1}^k m_i \quad (23)$$

$$\text{Calculate the total number of values known} \quad N = \sum_{i=1}^k n_i \quad (24)$$

$$\text{Calculate subgroup means} \quad \bar{x}_i = (1 - \mu_i) x_{i n_i} + \mu_i \sum_{j=1}^{n_i-1} \frac{x_{ij}}{(n_i - 1)} \quad (25)$$

$$\text{Calculate sub-group variances} \quad \sigma_i^2 = \alpha_i \sum_{j=1}^{n_i-1} (x_{in_i} - x_{ij})^2 + \beta_i \left[\sum_{j=1}^{n_i-1} (x_{in_i} - x_{ij}) \right]^2 \quad (26)$$

$$\text{Calculate z-mean} \quad \bar{z} = \frac{\sum_{i=1}^k n_i z_i}{N} \quad (27)$$

$$\text{Calculate group general mean} \quad \bar{\bar{x}} = \frac{\sum_{i=1}^k n_i \bar{x}_i}{N} \quad (28)$$

$$\text{Calculate mean variance factor} \quad \varepsilon = \frac{\sum_{i=1}^k \varepsilon_i}{k} \quad (29)$$

$$\text{Calculate residual variance} \quad \sigma_D^2 = \frac{\varepsilon \left[\sum_{i=1}^k \sigma_i^2 (n_i - 1) \right]}{(N - k)} \quad (30)$$

$$\text{Calculate x-sum of squares} \quad SSx = \sum_{i=1}^k n_i \bar{x}_i^2 - N \bar{\bar{x}}^2 \quad (31)$$

$$\text{Calculate z-sum of squares} \quad SSz = \sum_{i=1}^k n_i z_i^2 - N \bar{z}^2 \quad (32)$$

$$\text{Calculate xy-sum of products} \quad SPxz = \sum_{i=1}^k n_i \bar{x}_i z_i - N \bar{\bar{x}} \bar{z} \quad (33)$$

Calculate slope b
$$b = \frac{SP_{xz}}{SS_z} \quad (34)$$

Calculate intercept a
$$a = \bar{\bar{x}} - b \bar{z} \quad (35)$$

Calculate r^2
$$r^2 = \left(\frac{SP_{xz}}{SS_x} \right) \left(\frac{SP_{xz}}{SS_z} \right) \quad (36)$$

(r is the correlation coefficient)

Calculate non-linearity variance
$$\sigma_N^2 = \frac{(1 - r^2)SS_x}{(k - 2)} \quad (37)$$

Test equality of subgroup variances:

Calculate c
$$c = 1 + \frac{\left[\sum_{i=1}^k \frac{1}{(n_i - 1)} - \frac{1}{(N - k)} \right]}{3(k - 1)} \quad (38)$$

Calculate adjustment factor
$$A = 1 + \frac{\left(1 - \frac{N}{M} \right) \times \left(1 - \frac{12}{M} \right)}{2} \quad (39)$$

Calculate χ^2
$$\chi^2 = \frac{A}{c} \left[(N - k) \ln \left(\frac{\sigma_D^2}{\varepsilon} \right) - \sum_{i=1}^k (n_i - 1) \ln(\sigma_i^2) \right] \quad (40)$$

Test significance of deviations from linearity

Calculate F
$$F = \frac{\sigma_N^2}{\sigma_D^2} \quad (41)$$

Degrees of freedom for F $N - k$ (denominator), $k - 2$ (numerator)

Calculate non-regression variance
$$\sigma_T^2 = \frac{((N - k)\sigma_D^2 + (k - 2)\sigma_N^2)}{(N - 2)} \quad (42)$$

Calculate t
$$t = \bar{\bar{x}} \sqrt{\frac{N}{\sigma_T^2}} \quad (43)$$

Correction to t for censoring
$$A = \frac{\left(1 - \frac{N}{M} \right)}{\left(6.2 + \frac{N}{6.4} - \frac{(M - N)}{10.7} \right)} \quad (44)$$

4.3.4 Difference of means of 2 sub-groups

The t function for the difference of two sub-groups is calculated using Equation 45.

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{((n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2)}{(n_1 + n_2 - 2)} \times \left(\frac{\varepsilon_1}{n_1} + \frac{\varepsilon_2}{n_2} \right)}} \quad (45)$$

4.4 Complete (uncensored) data

If it is thought desirable, the properties of complete (uncensored) data sub-groups may be calculated using Equations 16 to 18, 19 and 45. The values of the coefficients are:

$$\text{Calculation of mean of sub-group} \quad \mu = 1 - \frac{1}{n} \quad (46)$$

$$\text{Calculation of variance of sub-group} \quad \alpha = \frac{1}{(n-1)}, \beta = \frac{1}{n(n-1)} \quad (47)$$

$$\text{Factor for calculation of variance of sub-group mean } \varepsilon = 1 \quad (48)$$

5 Random number generation

5.1 General

The generation of normally (Gaussian) distributed pseudo-random numbers is a 2-stage process. For each of these numbers two random numbers, uniformly distributed in the range 0 to 1, are required. These are then used to calculate a number which is a member of a normally distributed population.

Random numbers of both these types are subsequently referred to as "variates": "uniform variates" or "normal variates".

5.2 Uniform variates

For generation of the uniformly distributed numbers there are three possible routes:

- the "RND" function of most computer programming languages;
- the "linear congruential" (RAN1) calculation procedure;
- the "subtractive" (RAN 3) calculation procedure.

The generator to be used must satisfy the following criteria.

- The number of distinct values which can be generated must be large enough for the greatest number of values required for the Monte Carlo simulations to be only a small proportion of it. The largest Monte Carlo simulation is likely to require about 1 000 million values.
- There must be no internal correlations between variates within a sequence generated as a sample of a single random population for a Monte Carlo simulation.

The RND function of the computer programming language cannot generally be used because the number of distinct values available for the variates is usually quite small (below 1 million and often only 60 000). Relatively recent 32-bit software has a larger number of available values (about 16×10^6) but even this is not adequate for this work. The function appears usually to be based on the linear congruential method without "shuffling", so that the number of available values for normal variates is about one half of this number.

The linear congruential method has not been used for this work, because initial investigation showed that the "shuffling" procedure used to eliminate internal correlations is not always effective (see Graph Set 1 and Graph Set 6). For the purposes of the investigation, a computer function referred to as "RAN1" has been used, based on the linear congruential method, see Graph Set 6.

The "subtractive" procedure has been found completely satisfactory. The greatest number of distinct values obtainable for the uniformly distributed variates is limited by the maximum size of the "long" integer variable in most computer programming languages (2,147 483 647 $\times 10^9$). However, since these variates are used in pairs, the maximum number of values

of the resulting normal variates may be very much larger than this on account of the “shuffling” procedure included. The computer generating function referred to as “RAN3” has been used.

5.3 Normal variates

For the generation of normally distributed variates, the Box-Muller transformation is used:

$$y = \sqrt{(-2 \ln x_1)} \cos(2\pi x_2) \text{ or} \quad (49)$$

$$y = \sqrt{(-2 \ln x_1)} \sin(2\pi x_2) \quad (50)$$

where x_1 and x_2 are uniform variates and y is the required normal variate.

The values from these expressions are independently normally distributed, but they are related to each other. They may therefore be used to generate separate population samples, but may not be used together to generate a sample group from one population. For this reason, two uniform variates are required to generate a single normal one.

A modification of the Box-Muller transformation has been developed to avoid the use of the trigonometric functions $\sin$ and $\cos$, which can be slow. This modification replaces the values of x by derived variates within the unit circle at the origin.

$$v = (2x - 1) \text{ and} \quad (51)$$

$$r = \sqrt{(2x_1 - 1)^2 + (2x_2 - 1)^2} \quad (52)$$

The $\sin$ and $\cos$ functions may be replaced by the ratios v_1/r and v_2/r giving the two values for the variates

$$z_1 = v_1 / r * \sqrt{-2 \ln r} \quad \text{and} \quad z_2 = v_2 / r * \sqrt{-2 \ln r} \quad (53)$$

Again, these are independently normally distributed, but cannot be used together. A further disadvantage of this modification is that modern co-processors have much faster $\sin$ and $\cos$ functions, while the time taken for the tests to ascertain whether the points are within the unit circle, and whether there is a stored value is considerable, so that there is no time advantage in the use of the modification (see Graph Set 6, Graph Set 7 and Graph Set 8).

A very useful discussion of random number generation will be found in [2].

6 Validation procedures

6.1 Size of samples

Each sample of data groups for estimation of the distribution function of a statistical property (e.g. the F -ratio) may be derived from an (effectively) infinite population having the distribution to be estimated.

The confidence interval associated with the number of values in the sample at any probability level is determined by the size of the number. The statistical function relating the standard error of an estimated number of values to the total number of data groups in the sample and the expected number is the cumulative binomial function. The results in Annex A have been derived using the binomial function.

The binomial function itself may be described in general terms as follows (extract from [3]).

“Suppose an event occurs with a probability p per trial. Then the probability P of its occurring k or more times in N trials is termed a cumulative binomial probability, and is related to the incomplete beta function “ $I_x(a,b)$ ” as follows:

$$P = \sum_{j=k}^N \binom{N}{j} p^j (1-p)^{N-j} = I_p(k, N-k+1) \quad (54)$$

In this equation, N is the number of groups in the sample, P is the probability for the required confidence interval, p is the probability level in the required distribution function and k is the expected number at level p ($k = Np$).

When N and k are large ($k > \text{about } 100$) the distribution of the number found is very close to normal, with a standard deviation $\sqrt{Np(1-p)}$ and mean k .

In this work it was felt that at probability levels down to about 0,001 the upper and lower confidence intervals of the number of values should be not more than about 10 % of Np . To achieve this, a sample size of 10^6 is necessary (see Annex A).

Two alternative procedures (see 6.2) have been used. In both, a sample of 1 000 000 groups is generated, each giving a single value of the required function. These values may be arranged in an array in increasing order, and/or included in a histogram of 100 segments. The 1 000th value (for example) in this array will be the distribution value associated with a probability of 0,001. This probability can be considered to be accurate within 10 %. The probability of error greater than this is less than 1 %.

The data in the histograms were used to prepare the graphical representations of the distributions in Graph Set 2 to Graph Set 6.

6.2 Validation of random number generation

The objective of this work is to establish the distribution functions of the statistical properties of censored data groups as required by IEC 60216-3. These properties are:

- F (variance ratio)
- Bartlett's χ^2 (Variance equality test)
- “Student's” t

The basic technique used to ascertain the statistical distribution functions is to generate a data group and/or sub-groups for which the function values can be calculated. This is repeated for a large number of instances of the same group, and the sets of function values examined either graphically or numerically (or both), enabling the distribution functions to be established empirically.

The same procedure is also applied to the validation of the random number generation process by executing the calculations on uncensored data groups. The distribution functions in this case are known. If the distribution functions found are different from these, the conclusion is that the generation process is in error, i.e. that the random number generation process is not producing accurately normally distributed variates. The procedure is also applied to single sub-groups, determining the distributions of the sub-group mean, variance, variance of the mean, and t -ratio.

7 Outline of tasks

7.1 Optimum size of samples

See, 4.2, 6.1 and Annex A.

7.2 Uniformly and normally distributed random number generation

Test validity of possible generation methods for uniformly distributed data (RAN 1 and RAN 3) and conversion to normally distributed data by means of the properties of subgroups and groups. For conversion to normally distributed data, the modified Box-Muller method (see Equations 51 to 53) has not been used, since it offers no advantage and there have been some indications that it may not be always valid. Reference [3] is the main reference for these tasks. A highly detailed theoretical treatment will be found in [7].

7.3 Examine properties of single sub-groups

7.3.1 Correlation of mean and standard deviation (SD)

Mean and SD estimates are independent for complete (uncensored) data groups.

The correlation coefficients of mean and standard deviation (theoretically zero) are listed in Annex A.2 for complete and censored sub-groups.

7.3.2 Distribution and possible bias of mean and SD estimates

Determinations of mean, variance, standard error of mean are in Annex A.3. The distributions of estimates of t (ratio of mean to its standard error) are listed in Annex A.4

7.4 Examine properties of compound groups

The details of the properties to be examined are in 8.5. The details of the groups examined are in Annex A.6 and numerical details of some of the results obtained in Annex A.5.

7.5 Procedure for testing difference of two means

This is a widely used test, which is not at present directly used in IEC 60216. It is relevant to the calculations of IEC 60216-5 and 60216-6 for the relative thermal endurance characteristics.

The results are presented graphically in Graph Sets 7A and 7B.

8 Discussion of results and conclusions

8.1 Optimum size of samples

In order to achieve satisfactory precision of the final distributions, a sample size of 10^6 groups or sub-groups has generally been used.

8.2 Correlation coefficients of mean and standard deviation of sub-groups

For complete sub-groups of normally distributed variates, the correlation coefficient of mean and standard deviation should be zero, since theoretically they are independently distributed. This has been found to be the case (see Annex A.2, row 1). It is not true for censored data, and values for the regression constants and correlation coefficients are given in Annex A.2, rows 2 and 3.

Directly related to this is the finding that the ratio of mean to SD of mean for censored data sub-groups is no longer t -distributed. See 8.5.2 for details.

8.3 Values of mean, variance and SD of mean of censored data sub-groups

See Annex A.3 for tabulated values as above, both found and expected.

8.4 Distribution functions of mean and variance of censored data sub-groups

The distribution functions of some complete and some censored data sub-groups are shown graphically in the graphs of Graph Set 1.

The mean of censored sub-groups has been found to be normally distributed. However, the variance is not distributed according to the expected χ^2 -function having appropriate degrees of freedom.

8.5 Properties of compound data groups

Details of the groups analysed are to be found in Annex A.6. The group structure is essentially similar to the structure of the data from “one-way” analysis of variance problems.

A table of the numerical values of some group function distributions, censored and complete are in Annex A.5.

8.5.1 F (variance ratio) function

The values of the variance ratio function (F) are shown graphically in the graphs of Graph Set 2 (complete data) and Graph Set 3 and Graph Set 4 (censored). It is concluded that for F, the standard statistical function Tables may be employed as in one-way ANOVA, since the deviations from the expected distribution do not appear to be significant at any level of practical interest.

8.5.2 “Student’s” t function

8.5.2.1 Distribution

It has been found that the distribution function is related to the t -distribution function (t_0) by the following approximation:

$$1/t = 1/t_0 + A, \quad (55)$$

this being the equation of a rectangular hyperbola at 45° to the axes.

The value of the adjustment A is dependent on the size and degree of censoring of the data group and its constituent sub-groups.

The function for the calculation of A has been established by multi-stage linear regression methods, using variable A with values established at probability levels of 0,05 and 0,01. The initial regression parameters are r and n . The result is in Equation 56.

$$A = \frac{\left(1 - \frac{r}{n}\right)}{\left(6.2 + \frac{r}{6.4} - \frac{(n-r)}{10.7}\right)} \quad (56)$$

In the above equation, n and r are respectively the total number of data in a group or subgroup and the number of values known. The correction for a single sub-group is the same as for a composite group.

The value of the t -ratio is in reasonable agreement with the tabulated value after adjustment as indicated in Annex A.4 and Equation 55.

8.5.2.2 Application

The t function is generally used for one of two purposes:

- a) To establish whether the mean value obtained in a set of data is significantly different from an expected value. For this purpose, the t value is adjusted $1/t' = 1/t - A$ and compared with the tabulated value.
- b) To determine a confidence interval for the mean. For this purpose, the tabulated value for the appropriate number of degrees of freedom is adjusted $1/t = 1/t_0 + A$. The result is used in the equation for the calculation of confidence interval.

Graphs showing the two modes are in Graph Set 5.

8.5.3 Bartlett's χ^2 (variance equality) function

An interesting result is that the Bartlett's χ^2 -function for complete data is quite accurately in agreement with the tabulated χ^2 -function, while in the literature the relation is described as being "approximate" (see [4] and [5]).

For censored data the distribution is not the theoretical χ^2 .

The correction multiplier for the χ^2 function can be calculated as follows:

$$A = 1 + \frac{\left(1 - \frac{r}{n}\right) \times \left(1 - \frac{12}{n}\right)}{2} \quad (57)$$

where r and n are as in Equation 56.

Annex A.5 tabulates the distributions found for F , t and Bartlett's χ^2 with the multiplier factors found for χ^2 .

8.6 Generation of uniformly distributed random numbers

The group distribution functions for the F -function and χ^2 -function shown in Graph Set 6 (random generation function RAN1 {Linear congruential method} used as a first stage in generating normally distributed variates) exhibit substantial departures from normal. This is possibly due to internal correlations within the number population generated, and has led to the choice of the alternative (subtractive) generation method for preference.

NOTE This defect of RAN1 is only observed occasionally, and for very large groups.

8.7 Values of the t -ratio for the difference of two means

Graph Set 8 comprises graphs of the distribution functions of the ratio of the difference of two means to the standard error of this difference (t -ratio). Several different sub-group sizes and censoring degrees are included. The main conclusion is that provided the censoring degrees of the two sub-groups are approximately equal, the standard t -tables may be used without adjustment.

The correction for use of t -tables when necessary is similar to that of 8.5.2, the value of the correction being calculated as follows

$$A = \frac{p}{(r_1 + r_2)^2} \left(\frac{r_1}{n_1} - \frac{r_2}{n_2} \right) \left(2 + \left(\frac{r_1 + r_2}{20} \right)^2 \right) \quad (58)$$

where p is the smaller of r_1 and r_2 .

9 Conclusions

9.1 Random number generation

9.1.1 Uniform distribution

A frequent use of uniformly distributed random numbers is as the first stage of generation of groups of normally distributed random numbers. For very large data groups ($\geq$ about 100) the most widely used algorithm for computer generation of uniformly distributed pseudo-random numbers is the “linear congruential” (RAN 1). This appears to suffer intermittent defects which lead to non-normal distribution of the product.

The “subtractive” method of D.E.Knuth (RAN 3) has been found to be consistently satisfactory and has roughly the same speed of operation.

9.1.2 Normal distribution

For derivation of normally distributed variates from uniform, the simple Box-Muller relation is used, even though it requires two uniform for each normally distributed number. The modified calculation, which can produce two normal for each two uniform, and eliminates the use of sine and cosine functions (reputedly very slow) has not been found to have any speed advantage. Further, the two normal variates produced are not independent.

9.2 Statistical functions of censored data groups

9.2.1 Simple sub-groups

9.2.1.1 Group mean and variance

The group mean estimate has been found to be unbiased (accurate) and itself normally distributed.

The group variance estimate is unbiased, but does not have the expected distribution.

The variance estimate for the mean is again unbiased.

The mean and group standard deviation are correlated, which is not the case for uncensored data.

9.2.1.2 Confidence limits

The ratio of the mean to its standard deviation (the t function) is not distributed as the theoretical t -function, but a very simple adjustment dependent only on the group size and censored amount corrects quite closely. This correction enables the confidence limits of the mean estimate to be calculated using the tabulated t values (see Equations 55 and 56). (Adjustment of the tabulated t value is made by subtracting a from the reciprocal of the tabulated value).

9.2.2 Compound groups – Analysis of variance

Typical examples of compound groups, comprising more than one sub-group, are the data for one-way analysis of variance and the regression analysis of thermal endurance data in IEC 60216. For these, the most used statistical properties are

- variance equality: the χ^2 function testing the equality of the variances of the sub-groups;
- variance ratio: the F -function, testing whether the variance of the group of means of the sub-groups is compatible with the variances within the sub-groups;
- confidence limits of the general mean of all the data, calculated using the t -function.

9.2.2.1 Variance equality

The χ^2 function has not been found to conform to the χ^2 distribution, but this is remedied by a simple multiplier correction, whose value is found from the data numbers of the subgroups (see Equation 57).

9.2.2.2 Variance ratio

The F -function has been found to be F -distributed, and its significance can be assessed in the usual way.

9.2.2.3 Confidence limits of the general mean

The t -function is found to have the same distribution properties for compound groups as for single sub-groups, and the correction is calculated and applied in the same way.

9.2.3 Compound groups – Regression analysis

9.2.3.1 Slope and intercept estimates

The inaccuracies of the means of the slope and intercept have all been less than 10^{-3} times the parent data standard deviation (equal to 1 in the test samples). The standard deviation of the 10^6 means in all cases so far has been about 0,2 and that of the intercepts 0,4 to 0,5.

9.2.3.2 Variance ratio

The distribution of F in the great majority of cases is acceptably close to theoretical. Only in an unusual case of widely different sub-group size and fraction censored (6/11, 11/21 and 21/21) has a result of doubtful acceptability been found.

For the more usual sub-group dimensions good agreement has been found. These cases include:

- equal size and censoring (e.g. 11/21, 11/21, 11/21);
- equal size and variable censoring (e.g. 11/21, 15/21, 21/21);
- variable size and equal uncensored remainder (e.g. 11/21, 11/18, 11/11);
- group sizes greater than about 30.

9.2.3.3 Variance equality (χ^2)

The variance equality function (χ^2) has the same distribution function as for analysis of variance, and therefore has the same correction function.

9.2.3.4 t -function (for confidence estimates)

There is again the same behaviour as for analysis of variance.

9.2.4 Difference of means of 2 sub-groups

This involves the t -function, calculated using a slight modification of the method for complete data (see Equation 45 compared with Equation 15). The correction is given by Equation 58 applied in the same way as for compound groups (Equation 55).

If the degree of censoring is the same for both groups, even though they are of different sizes, the distribution of the t -function estimate is the theoretical. The difference from theory is not significant provided the censoring is approximately the same for both.

10 Summary of conclusions and recommendations

10.1 Summary of conclusions

- Estimates of primary statistical functions (mean, variance and variance of mean of a censored data sub-group) are unbiased. The distribution of the mean is normal, but the distributions of the other two are not as for complete data.
- Estimates of secondary statistical functions (F , t and χ^2) are quite close to the functions of complete data if simple corrections are made to t and χ^2 .
- Uniformly distributed random numbers are best generated using the “subtractive” method.
- The “linear congruential” method is not satisfactory for generation of very large samples (more than about 1 million uniformly distributed numbers).
- For comparison of two means, if the data groups have approximately the same censoring fraction, the t function is almost exactly theoretically distributed. If the groups are the same in size and censoring point, the distribution is exact.

10.2 Recommendations

- Censoring before the median value is not provided for in these calculations.
- For single “sub-group” tests (comparison of the mean with a constant) ensure that the number of data values is at least 11.
- For composite group tests (e.g. ANOVA or regression analysis) design the sub-groups to have as many properties the same as possible e.g.
 - same size, variable censoring (the most likely case for aging tests to IEC 60216);
 - variable size, approximately same censoring fraction;
 - variable size, same known number of data.
- For comparison of two means, censor after the same fraction.

Annex A (informative)

Results

A.1 Confidence limits for value counts

(Results derived using the binomial function)

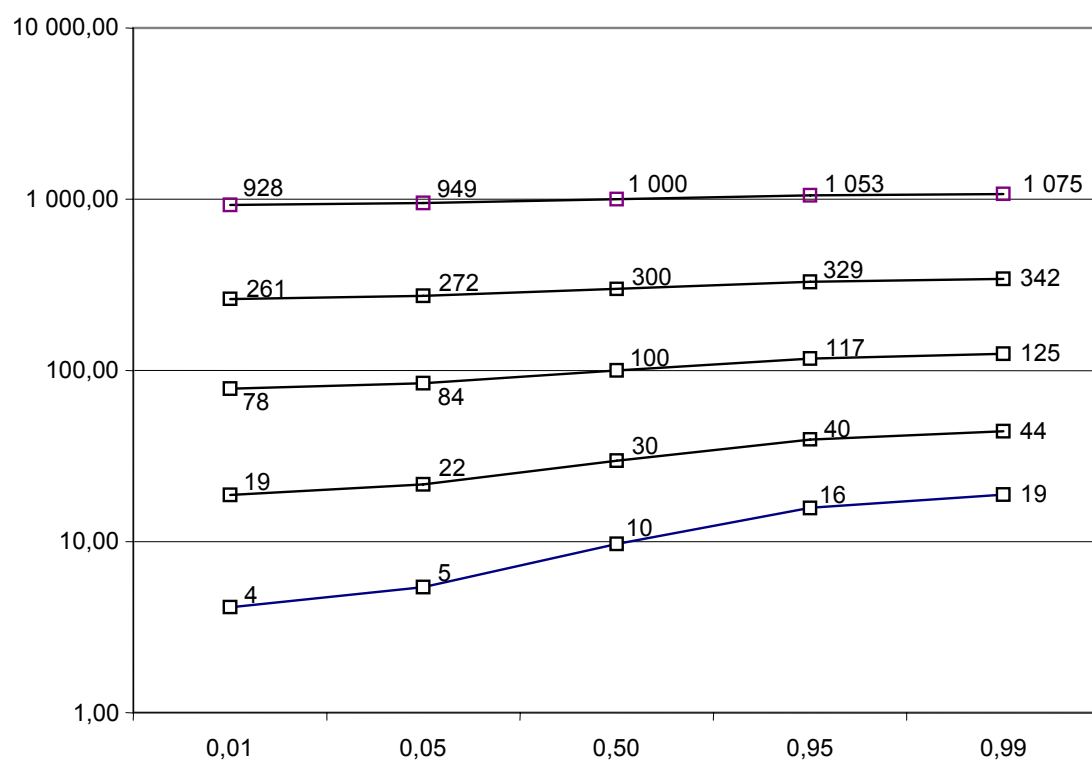
Table A.1 below shows the confidence limits for the numbers of function values to be found in a data set of 1 000 000 ($N=10^6$) groups

Table A.1 – Confidence limits for value counts

Probability of value, P	Median rank	Upper 1% limit	Upper 5% limit	50% (median rank)	Lower 5% limit	Lower 1% limit
0,00001	10	19	16	10	5	4
0,00003	30	44	40	30	22	19
0,0001	100	145	117	100	84	78
0,0003	300	342	329	300	272	261
0,001	1000	1075	1053	1000	949	928

If the probability of a function value is P , the median rank value of the number of times this value is found is NP . The upper 1 % confidence limit of this is in column 3 above, and similarly for the other confidence limits.

For $P>0,001$, the distribution is for all purposes normal, and the standard deviation is $\sqrt{NP(1-P)}$.



IEC 551/10

Figure A.1 – Confidence limits for value counts

A.2 Correlation of mean and standard deviation

The table below shows the regression and correlation of standard deviation on mean for censored and uncensored data sub-groups.

The first set of values is for complete data. For this case the expected value of the correlation coefficient is zero.

Values are based on data sets of 10⁶ single sub-groups

Group size	5	5	5	5	5	10	20	50	100										
Known	5	5	5	5	5	10	20	50	100										
slope	-6,13E-04	-1,59E-03	-1,63E-04	2,94E-04	-1,19E-04	-1,06E-03	-1,38E-03												
intercept	0,940189	0,940122	0,939803	0,973246	0,988822	0,994722	0,99747												
correlation	-8,04E-04	-2,08E-03	-2,14E-04	4,00E-04	-1,65E-04	-1,48E-03	-1,94E-03												
Group size	5	5	5	5	5	6	7	11	21	21	21	21	21	21	21	21	21	21	21
Known	3	3	3	3	3	4	4	6	11	11	11	11	11	11	11	11	11	11	11
slope	0,370926	0,370909	0,372341	0,269255	0,378173	0,380799	0,382693	0,383733	0,383961										
intercept	0,868381	0,868734	0,868718	0,911668	0,907774	0,941961	0,970171	0,97016	0,970213										
correlation	0,406281	0,406487	0,407512	0,299238	0,415865	0,42219	0,427132	0,428162	0,428729										
Group size	31	41	41	41	41	41	101												
Known	18	21	28	35	53														
slope	0,311667	0,384046	0,201394	6,75E-02	0,362531														
intercept	0,982404	0,985144	0,989413	0,992069	0,994212														
correlation	0,351275	0,43055	0,235212	8,56E-02	0,408956														

A.3 Averages of 1 000 000 (10⁶) groups

Group	Known	Mean (Expected = 0)	Variance (Expected = 1)	Found	SD of Mean	Expected
5	3	0,000608	0,994104	0,536516		0,535552
5	4	-0,000599	1,000401	0,467841		0,46746
7	4	-0,000297	0,99745	0,458515		0,458742
7	6	0,000094	1,00019	0,387176		0,38704
11	6	-0,000845	0,996903	0,370226		0,370354
11	7	-0,00012	0,999611	0,340764		0,34072
11	9	-0,000239	0,999037	0,311524		0,311508
11	10	0,00033	1,000473	0,30539		0,304974
21	11	-0,000351	0,999554	0,270723		0,270666
21	14	0,000169	1,000394	0,240231		0,240054
21	17	0,000204	0,999748	0,225488		0,22559
21	20	0,000432	1,004482	0,219442		0,219221
31	16	0,00024	1,0003	0,22362		0,223531
31	25	-0,0001	0,999599	0,185909		0,185709
31	28	0,000016	1,001471	0,181488		0,18164
31	30	0,000288	1,006427	0,179994		0,180108
41	21	-0,000307	0,999642	0,194962		0,194702
41	32	-0,000229	1,000169	0,162711		0,162748
41	35	-0,000011	0,999444	0,15938		0,159452
41	38	0,000405	1,003532	0,157255		0,157356
51	26	0,000385	1,000705	0,175206		0,174757
51	36	-0,000428	0,999701	0,15037		0,150343
51	46	0,000014	1,001497	0,141665		0,141609

A.4 t-values for single sub-groups

n = total number t_{th} = theoretical (tabulated) value of t

$$1/t_a = 1/t_{th} + a$$

r = number known t_a = adjusted value of t

$$a = \frac{\left(1 - \frac{r}{n}\right)}{\left(6.2 + \frac{r}{6.4} - \frac{(n-r)}{10.7}\right)}$$

Observed values										Adjusted theoretical values										
n	r	df	a	0,999	0,995	0,99	0,95	0,5	0,05	0,01	0,005	0,001	0,999	0,995	0,99	0,95	0,05	0,01	0,005	0,001
5	3	2	0,06171	-43,59	-17,89	-12,29	-4,65	0	2	4,16	5,72	11,8	59,08	-25,61	-12,21	-3,56	2,47	4,87	6,16	9,39
5	4	3	0,02971	-13,4	-7,38	-5,63	-2,76	0	2,16	3,99	5,08	8,75	-14,66	-7,07	-5,25	-2,53	2,20	4,00	4,98	7,84
6	4	3	0,05022	-16,23	-8,77	-6,62	-3,1	0	1,94	3,43	4,3	7,28	-20,97	-8,26	-5,88	-2,67	2,10	3,70	4,52	6,75
7	4	3	0,06548	-18,18	-9,79	-7,31	-3,39	0	1,73	2,93	3,61	5,81	-30,85	-9,46	-6,46	-2,78	2,04	3,50	4,22	6,12
6	5	4	0,02420	-8,58	-5,35	-4,32	-2,36	0	2,01	3,46	4,24	6,5	-8,68	-5,18	-4,12	-2,25	2,03	3,44	4,14	6,11
7	5	4	0,04205	-10,06	-6,13	-4,88	-2,58	0	1,87	3,11	3,79	5,83	-10,27	-5,71	-4,45	-2,34	1,96	3,24	3,86	5,51
8	5	4	0,05596	-11,24	-6,8	-5,35	-2,76	0	1,75	2,8	3,36	5,04	-11,98	-6,20	-4,74	-2,42	1,90	3,10	3,66	5,12
9	5	4	0,06726	-11,84	-7,24	-5,68	-2,91	0	1,64	2,56	3,02	4,34	-13,86	-6,67	-5,01	-2,49	1,86	2,99	3,52	4,84
7	6	5	0,02028	-6,77	-4,53	-3,73	-2,17	0	1,94	3,19	3,82	5,62	-6,69	-4,39	-3,61	-2,10	1,94	3,15	3,73	5,26
8	6	5	0,03597	-7,49	-4,98	-4,1	-2,32	0	1,84	2,95	3,49	5,04	-7,48	-4,72	-3,83	-2,17	1,88	3,00	3,52	4,86
9	6	5	0,04861	-8,38	-5,44	-4,43	-2,44	0	1,74	2,74	3,21	4,46	-8,26	-5,02	-4,02	-2,23	1,84	2,89	3,37	4,58
10	6	5	0,05914	-8,9	-5,79	-4,66	-2,56	0	1,66	2,55	2,97	4,14	-9,05	-5,29	-4,20	-2,29	1,80	2,81	3,26	4,37
11	6	5	0,06815	-9,44	-6,08	-4,92	-2,67	0	1,61	2,41	2,77	3,77	-9,85	-5,56	-4,37	-2,34	1,77	2,74	3,16	4,20
8	7	6	0,01736	-5,83	-4,09	-3,42	-2,06	0	1,88	3,01	3,53	4,95	-5,73	-3,96	-3,32	-2,01	1,88	2,98	3,48	4,78
9	7	6	0,03127	-6,42	-4,38	-3,65	-2,16	0	1,81	2,85	3,33	4,62	-6,22	-4,19	-3,49	-2,07	1,83	2,86	3,32	4,48
10	7	6	0,04278	-6,87	-4,68	-3,88	-2,26	0	1,74	2,7	3,14	4,3	-6,70	-4,41	-3,63	-2,12	1,79	2,77	3,20	4,26
11	7	6	0,05255	-7,27	-4,93	-4,07	-2,35	0	1,67	2,55	2,95	3,95	-7,17	-4,60	-3,76	-2,16	1,76	2,70	3,10	4,09
12	7	6	0,06104	-7,75	-5,21	-4,29	-2,44	0	1,61	2,41	2,76	3,69	-7,63	-4,79	-3,89	-2,20	1,74	2,64	3,02	3,95
13	7	6	0,06855	-8,05	-5,38	-4,42	-2,5	0	1,57	2,29	2,61	3,41	-8,10	-4,97	-4,01	-2,24	1,71	2,59	2,96	3,84
9	8	7	0,01510	-5,28	-3,77	-3,2	-1,98	0	1,85	2,9	3,38	4,61	-5,16	-3,69	-3,14	-1,95	1,84	2,87	3,32	4,46
10	8	7	0,02754	-5,7	-4,05	-3,4	-2,06	0	1,8	2,79	3,24	4,39	-5,51	-3,87	-3,27	-2,00	1,80	2,77	3,19	4,23
11	8	7	0,03804	-6	-4,23	-3,56	-2,14	0	1,74	2,65	3,06	4,12	-5,85	-4,04	-3,38	-2,04	1,77	2,69	3,09	4,05
12	8	7	0,04711	-6,26	-4,47	-3,73	-2,21	0	1,69	2,54	2,93	3,88	-6,18	-4,19	-3,49	-2,08	1,74	2,63	3,00	3,91
13	8	7	0,05508	-6,63	-4,64	-3,88	-2,28	0	1,63	2,43	2,78	3,64	-6,50	-4,34	-3,59	-2,12	1,72	2,57	2,93	3,79
8	8	7	0,06221	-6,96	-4,82	-4,01	-2,33	0	1,59	2,32	2,64	3,41	-6,81	-4,47	-3,69	-2,15	1,69	2,53	2,87	3,69
9	8	7	0,06867	-7,2	-4,95	-4,1	-2,38	0	1,55	2,24	2,52	3,25	-7,13	-4,61	-3,78	-2,18	1,68	2,49	2,82	3,60

				Observed values					Amended theoretical values											
<i>n</i>	<i>r</i>	<i>df</i>	<i>a</i>	0,999	0,995	0,99	0,95	0,5	0,05	0,01	0,005	0,001	0,999	0,995	0,99	0,95	0,05	0,01	0,005	0,001
10	9	8	0,01331	-4,86	-3,57	-3,06	-1,93	0	1,82	2,82	3,25	4,33	-4,79	-3,51	-3,01	-1,91	1,81	2,79	3,21	4,25
11	9	8	0,02451	-5,18	-3,78	-3,21	-2	0	1,77	2,71	3,13	4,17	-5,06	-3,66	-3,12	-1,95	1,78	2,70	3,10	4,05
12	9	8	0,03413	-5,47	-3,95	-3,34	-2,05	0	1,73	2,62	3	3,96	-5,32	-3,79	-3,21	-1,99	1,75	2,64	3,01	3,90
13	9	8	0,04254	-5,73	-4,12	-3,48	-2,12	0	1,69	2,53	2,89	3,8	-5,57	-3,91	-3,30	-2,02	1,72	2,58	2,94	3,78
14	9	8	0,05003	-5,94	-4,26	-3,59	-2,17	0	1,64	2,43	2,77	3,56	-5,81	-4,03	-3,39	-2,05	1,70	2,53	2,87	3,67
15	9	8	0,05677	-6,22	-4,39	-3,7	-2,22	0	1,6	2,35	2,66	3,41	-6,05	-4,14	-3,47	-2,08	1,68	2,49	2,82	3,58
16	9	8	0,06293	-6,44	-4,53	-3,8	-2,26	0	1,57	2,27	2,55	3,24	-6,28	-4,25	-3,54	-2,11	1,66	2,45	2,77	3,51
17	9	8	0,06861	-6,59	-4,65	-3,9	-2,3	0	1,54	2,2	2,47	3,11	-6,51	-4,36	-3,61	-2,13	1,65	2,42	2,73	3,44
15	10	9	0,04569	-5,55	-4,01	-3,41	-2,09	0	1,66	2,43	2,76	3,55	-5,35	-3,82	-3,24	-2,00	1,69	2,50	2,83	3,59
16	10	9	0,05207	-5,67	-4,11	-3,49	-2,13	0	1,62	2,36	2,67	3,43	-5,54	-3,91	-3,31	-2,03	1,67	2,46	2,78	3,51
17	10	9	0,05793	-5,87	-4,23	-3,57	-2,17	0	1,59	2,3	2,59	3,27	-5,72	-4,00	-3,37	-2,05	1,66	2,43	2,73	3,44
18	10	9	0,06336	-6,03	-4,34	-3,66	-2,21	0	1,56	2,23	2,51	3,14	-5,90	-4,09	-3,44	-2,07	1,64	2,39	2,69	3,38
19	10	9	0,06844	-6,13	-4,42	-3,73	-2,24	0	1,53	2,18	2,44	3,05	-6,09	-4,18	-3,50	-2,10	1,63	2,36	2,66	3,32
12	11	10	0,01065	-4,36	-3,31	-2,87	-1,86	0	1,78	2,71	3,09	4,06	-4,33	-3,28	-2,85	-1,85	1,78	2,68	3,07	3,97
16	11	10	0,04194	-5,15	-3,81	-3,26	-2,03	0	1,66	2,44	2,76	3,52	-5,02	-3,66	-3,13	-1,96	1,68	2,48	2,80	3,53
17	11	10	0,04797	-5,34	-3,91	-3,35	-2,07	0	1,63	2,38	2,7	3,4	-5,17	-3,74	-3,19	-1,99	1,67	2,44	2,75	3,46
18	11	10	0,05353	-5,45	-3,97	-3,4	-2,1	0	1,6	2,32	2,61	3,28	-5,32	-3,82	-3,24	-2,01	1,65	2,41	2,71	3,39
19	11	10	0,05872	-5,61	-4,09	-3,48	-2,13	0	1,58	2,26	2,54	3,18	-5,48	-3,89	-3,30	-2,03	1,64	2,38	2,67	3,33
20	11	10	0,06358	-5,66	-4,16	-3,53	-2,16	0	1,55	2,2	2,46	3,06	-5,63	-3,97	-3,35	-2,05	1,63	2,35	2,64	3,28
2	11	10	0,06818	-5,82	-4,2	-3,58	-2,19	0	1,53	2,17	2,42	3	-5,78	-4,04	-3,41	-2,07	1,61	2,33	2,61	3,23
18	12	11	0,04436	-5,03	-3,72	-3,2	-2,02	0	1,64	2,39	2,69	3,4	-4,90	-3,60	-3,09	-1,95	1,66	2,43	2,73	3,42
19	12	11	0,04965	-5,17	-3,82	-3,27	-2,05	0	1,62	2,34	2,62	3,29	-5,03	-3,67	-3,14	-1,97	1,65	2,39	2,69	3,35
20	12	11	0,05459	-5,3	-3,89	-3,33	-2,08	0	1,59	2,28	2,56	3,21	-5,16	-3,74	-3,19	-1,99	1,64	2,37	2,66	3,30
21	12	11	0,05925	-5,41	-3,97	-3,4	-2,1	0	1,56	2,24	2,5	3,1	-5,28	-3,81	-3,24	-2,01	1,62	2,34	2,62	3,25
22	12	11	0,06366	-5,48	-4,01	-3,43	-2,13	0	1,54	2,19	2,44	3	-5,41	-3,87	-3,29	-2,03	1,61	2,32	2,59	3,20
23	12	11	0,06787	-5,61	-4,09	-3,49	-2,15	0	1,53	2,15	2,39	2,93	-5,54	-3,94	-3,33	-2,05	1,60	2,29	2,57	3,16
20	13	12	0,04619	-4,92	-3,68	-3,17	-2	0	1,62	2,34	2,62	3,29	-4,80	-3,56	-3,06	-1,94	1,65	2,39	2,68	3,33
21	13	12	0,05091	-4,97	-3,74	-3,22	-2,03	0	1,6	2,3	2,58	3,22	-4,91	-3,62	-3,10	-1,96	1,63	2,36	2,64	3,27
22	13	12	0,05536	-5,11	-3,82	-3,28	-2,05	0	1,58	2,25	2,51	3,12	-5,02	-3,68	-3,15	-1,98	1,62	2,33	2,61	3,23
23	13	12	0,05959	-5,2	-3,86	-3,32	-2,08	0	1,56	2,22	2,47	3,04	-5,13	-3,73	-3,19	-1,99	1,61	2,31	2,58	3,18
24	13	12	0,06363	-5,32	-3,93	-3,37	-2,1	0	1,54	2,18	2,42	2,97	-5,24	-3,79	-3,23	-2,01	1,60	2,29	2,56	3,14
25	13	12	0,06751	-5,33	-3,96	-3,4	-2,12	0	1,53	2,14	2,37	2,91	-5,35	-3,85	-3,27	-2,03	1,59	2,27	2,53	3,11
22	14	13	0,04760	-4,81	-3,63	-3,13	-1,99	0	1,61	2,31	2,59	3,22	-4,72	-3,52	-3,03	-1,93	1,63	2,35	2,63	3,26
23	14	13	0,05185	-4,9	-3,69	-3,18	-2,01	0	1,59	2,26	2,53	3,13	-4,81	-3,57	-3,07	-1,95	1,62	2,33	2,61	3,21
24	14	13	0,05591	-4,97	-3,74	-3,22	-2,04	0	1,57	2,23	2,49	3,06	-4,91	-3,62	-3,11	-1,97	1,61	2,31	2,58	3,17
25	14	13	0,05979	-5,06	-3,79	-3,26	-2,05	0	1,55	2,2	2,44	3,01	-5,00	-3,67	-3,15	-1,98	1,60	2,29	2,55	3,13
26	14	13	0,06352	-5,15	-3,86	-3,32	-2,08	0	1,54	2,16	2,4	2,94	-5,10	-3,73	-3,19	-2,00	1,59	2,27	2,53	3,09
27	14	13	0,06713	-5,25	-3,9	-3,35	-2,09	0	1,52	2,14	2,37	2,88	-5,20	-3,78	-3,22	-2,01	1,58	2,25	2,51	3,06

Observed values										Amended theoretical values										
<i>n</i>	<i>r</i>	<i>df</i>	<i>a</i>	0,999	0,995	0,99	0,95	0,5	0,05	0,01	0,005	0,001	0,999	0,995	0,99	0,95	0,05	0,01	0,005	0,001
24	15	14	0,04868	-4,72	-3,56	-3,09	-1,98	0	1,6	2,29	2,56	3,16	-4,64	-3,48	-3,01	-1,93	1,62	2,33	2,60	3,20
25	15	14	0,05257	-4,78	-3,62	-3,13	-1,99	0	1,58	2,24	2,5	3,08	-4,73	-3,53	-3,04	-1,94	1,61	2,31	2,57	3,16
26	15	14	0,05629	-4,9	-3,67	-3,17	-2,02	0	1,57	2,22	2,46	3,01	-4,81	-3,58	-3,08	-1,96	1,60	2,29	2,55	3,12
27	15	14	0,05988	-4,93	-3,7	-3,2	-2,03	0	1,55	2,19	2,44	2,99	-4,90	-3,62	-3,11	-1,97	1,59	2,27	2,53	3,09
28	15	14	0,06335	-5,05	-3,76	-3,24	-2,05	0	1,54	2,15	2,39	2,9	-4,98	-3,67	-3,15	-1,98	1,58	2,25	2,50	3,05
29	15	14	0,06672	-5,06	-3,79	-3,27	-2,07	0	1,53	2,13	2,35	2,86	-5,07	-3,71	-3,18	-2,00	1,58	2,23	2,48	3,02
26	16	15	0,04953	-4,62	-3,53	-3,06	-1,97	0	1,59	2,26	2,52	3,09	-4,58	-3,45	-2,99	-1,92	1,61	2,31	2,57	3,15
27	16	15	0,05310	-4,72	-3,57	-3,1	-1,98	0	1,57	2,23	2,48	3,03	-4,66	-3,49	-3,02	-1,93	1,60	2,29	2,55	3,12
28	16	15	0,05655	-4,78	-3,62	-3,13	-2	0	1,56	2,2	2,45	2,99	-4,73	-3,54	-3,05	-1,95	1,59	2,27	2,53	3,08
29	16	15	0,05989	-4,87	-3,67	-3,17	-2,02	0	1,55	2,18	2,42	2,95	-4,81	-3,58	-3,08	-1,96	1,59	2,25	2,50	3,05
30	16	15	0,06313	-4,91	-3,7	-3,2	-2,03	0	1,54	2,16	2,39	2,91	-4,88	-3,62	-3,11	-1,97	1,58	2,24	2,48	3,02
31	16	15	0,06630	-4,98	-3,73	-3,22	-2,04	0	1,53	2,14	2,37	2,85	-4,96	-3,66	-3,15	-1,98	1,57	2,22	2,47	2,99
28	17	16	0,05018	-4,57	-3,48	-3,03	-1,96	0	1,58	2,25	2,5	3,06	-4,52	-3,42	-2,97	-1,91	1,61	2,29	2,55	3,11
29	17	16	0,05350	-4,6	-3,54	-3,07	-1,97	0	1,57	2,23	2,47	3,02	-4,59	-3,46	-3,00	-1,93	1,60	2,27	2,53	3,08
30	17	16	0,05671	-4,71	-3,58	-3,1	-1,99	0	1,56	2,2	2,44	2,97	-4,66	-3,50	-3,03	-1,94	1,59	2,25	2,51	3,05
31	17	16	0,05983	-4,77	-3,61	-3,14	-2,01	0	1,54	2,17	2,41	2,93	-4,73	-3,54	-3,06	-1,95	1,58	2,24	2,49	3,02
32	17	16	0,06288	-4,84	-3,65	-3,16	-2,02	0	1,54	2,15	2,39	2,9	-4,80	-3,58	-3,08	-1,96	1,57	2,22	2,47	2,99
30	18	17	0,05069	-4,57	-3,48	-3,02	-1,95	0	1,58	2,23	2,48	3,01	-4,47	-3,40	-2,95	-1,91	1,60	2,27	2,53	3,08
31	18	17	0,05378	-4,58	-3,5	-3,04	-1,96	0	1,57	2,21	2,46	2,99	-4,53	-3,43	-2,98	-1,92	1,59	2,26	2,51	3,05
32	18	17	0,05679	-4,64	-3,54	-3,07	-1,98	0	1,56	2,19	2,43	2,95	-4,60	-3,47	-3,00	-1,93	1,58	2,24	2,49	3,02
32	19	18	0,05108	-4,5	-3,44	-2,98	-1,94	0	1,58	2,23	2,48	3	-4,43	-3,37	-2,94	-1,90	1,59	2,26	2,51	3,05

For composite groups made up of several subgroups, the same relationship has been found to be applicable. The total number of data is the total of all subgroups, and the number of data known is the sum of the known numbers in the sub-groups.

A.5 Distribution Values for F , t and χ^2

The values in the following Table are based on sets of 10^6 groups. In generation of the groups, the values are saved in arrays, which are sorted when complete into increasing order. The value of the property whose sequence number is equal to the product of the probability (column headed p) and 10^6 is taken as the best estimate of the expected value of the property at this probability level.

The 2nd to 4th columns are the values calculated from the theoretical distributions. The 6th to 8th are the values from the sets generated. Column 9 is the ratio of the observed value of χ^2 to the theoretical value, and column 10 contains the value (A) of this ratio derived from a large number of data sets of varying sizes (Equation (35) in the main text):

$$A = 1 + \frac{\left(1 - \frac{r}{n}\right) \times \left(1 - \frac{12}{n}\right)}{2}.$$

3x5-3	9,99201E-05	-4,984904	1,99472E-04	0,0001	1,1859 E-4	-9,258230	2,65230E-04	1,329660	1,040000	9	15
	5,00044E-04	-4,140311	9,99722E-04	0,0005	5,7178 E-4	-6,855920	1,23100E-03	1,231342			
	1,00053E-03	-3,787427	2,00138E-03	0,001	1,1301 E-3	-6,035130	2,34220E-03	1,170295			
	5,01466E-03	-2,976849	1,00247E-02	0,005	5,3842 E-3	-4,323490	1,19770E-02	1,194753			
	1,00586E-02	-2,624492	2,01004E-02	0,01	1,0758 E-2	-3,661780	2,38630E-02	1,187190			
	5,15135E-02	-1,761309	1,02586E-01	0,05	5,533 E-2	-2,208500	1,22600E-01	1,195092			
	1,06291E-01	-1,345031	2,10721E-01	0,1	1,1437 E-1	-1,604910	2,52220E-01	1,196939			
	7,34772E-01	0,000000	1,38629E+00	0,5	8,0976 E-1	-0,002550	1,60790E+00	1,159855			
	2,80679E+00	1,345031	4,60518E+00	0,9	3,46390E+00	1,249180	5,02190E+00	1,090490			
	3,88529E+00	1,761309	5,99148E+00	0,95	5,13840E+00	1,604250	6,41810E+00	1,071205			
	6,92660E+00	2,624492	9,21035E+00	0,99	1,1063 E1	2,334210	9,51690E+00	1,033283			
	8,50969E+00	2,976849	1,05965E+01	0,995	1,4866 E1	2,640660	1,08160E+01	1,020712			
	1,29730E+01	3,787427	1,38150E+01	0,999	2,8617 E1	3,359470	1,37120E+01	0,992544			
	1,52977E+01	4,140311	1,52014E+01	0,9995	3,7519 E1	3,696360	1,49560E+01	0,983860			
	2,18570E+01	4,984904	1,84247E+01	0,9999	6,8701 E1	4,529200	1,78800E+01	0,970434			
3x21-11	9,99201E-05	-4,197937	1,99472E-04	0,0001	1,07220E-04	-5,301650	2,67090E-04	1,338984	1,192744	33	63
	5,00044E-04	-3,621826	9,99722E-04	0,0005	4,88000E-04	-4,473770	1,23280E-03	1,233143			
	1,00053E-03	-3,365276	2,00138E-03	0,001	1,02370E-03	-4,107520	2,37450E-03	1,186434			
	5,01332E-03	-2,738489	1,00247E-02	0,005	5,09230E-03	-3,222030	1,22450E-02	1,221487			
	1,00537E-02	-2,448678	2,01004E-02	0,01	1,02570E-02	-2,823590	2,46320E-02	1,225447			
	5,13811E-02	-1,693888	1,02586E-01	0,05	5,19410E-02	-1,868420	1,26210E-01	1,230282			
	1,05731E-01	-1,308573	2,10721E-01	0,1	1,07050E-01	-1,412360	2,58940E-01	1,228830			
	7,09412E-01	0,000000	1,38629E+00	0,5	7,12470E-01	0,001960	1,69830E+00	1,225065			
	2,48872E+00	1,308573	4,60518E+00	0,9	2,48540E+00	1,223150	5,56280E+00	1,207945			
	3,31583E+00	1,693888	5,99148E+00	0,95	3,30440E+00	1,548900	7,21640E+00	1,204444			
	5,39035E+00	2,448678	9,21035E+00	0,99	5,32600E+00	2,158330	1,10140E+01	1,195828			
	6,35464E+00	2,738489	1,05965E+01	0,995	6,26160E+00	2,381660	1,25720E+01	1,186426			
	8,77299E+00	3,365276	1,38150E+01	0,999	8,61580E+00	2,845680	1,62910E+01	1,179225			
	9,89712E+00	3,621826	1,52014E+01	0,9995	9,73930E+00	3,024360	1,78160E+01	1,172001			
	1,27184E+01	4,197937	1,84247E+01	0,9999	1,23850E+01	3,404480	2,13760E+01	1,160179			
3x21-11	9,99201E-05	-4,197937	1,99472E-04	0,0001	8,49512E-05	-5,455205	2,38704E-04	1,196679	1,192744	33	63
	5,00044E-04	-3,621826	9,99722E-04	0,0005	4,80640E-04	-4,473283	1,18557E-03	1,185898			
	1,00053E-03	-3,365276	2,00138E-03	0,001	9,81484E-04	-4,107124	2,43776E-03	1,218044			
	5,01332E-03	-2,738489	1,00247E-02	0,005	5,13827E-03	-3,231887	1,25426E-02	1,251175			
	1,00537E-02	-2,448678	2,01004E-02	0,01	1,01906E-02	-2,837668	2,49673E-02	1,242130			
	5,13811E-02	-1,693888	1,02586E-01	0,05	5,20794E-02	-1,869131	1,26515E-01	1,233258			
	1,05731E-01	-1,308573	2,10721E-01	0,1	1,06734E-01	-1,410576	2,59810E-01	1,232957			
	7,09412E-01	0,000000	1,38629E+00	0,5	7,12450E-01	0,000562	1,69469E+00	1,222461			
	2,48872E+00	1,308573	4,60518E+00	0,9	2,48334E+00	1,220718	5,55943E+00	1,207212			
	3,31583E+00	1,693888	5,99148E+00	0,95	3,29800E+00	1,552454	7,18802E+00	1,199708			
	5,39035E+00	2,448678	9,21035E+00	0,99	5,33357E+00	2,161683	1,09615E+01	1,190128			
	6,35464E+00	2,738489	1,05965E+01	0,995	6,29254E+00	2,388336	1,25694E+01	1,186185			
	8,77299E+00	3,365276	1,38150E+01	0,999	8,61813E+00	2,863686	1,62120E+01	1,173510			
	9,89712E+00	3,621826	1,52014E+01	0,9995	9,73647E+00	3,055030	1,77361E+01	1,166743			
	1,27184E+01	4,197937	1,84247E+01	0,9999	1,23981E+01	3,466062	2,16004E+01	1,172360			



Docu

ment

	Normal distribution			p	F	Censored normal			χ^2 ratio	χ^2 ratio calc	known	total
	F	t	χ^2			t	χ^2					
3x21-41	9,99201E-05	-3,953464	1,99472E-04	0,0001	1,08134E-04	-4,696092	2,63336E-04	1,320165	1,220107	63	123	
	5,00044E-04	-3,454479	9,99722E-04	0,0005	4,89965E-04	-4,035096	1,24682E-03	1,247171				
	1,00053E-03	-3,226960	2,00138E-03	0,001	9,97585E-04	-3,733343	2,43621E-03	1,217268				
	5,01288E-03	-2,657471	1,00247E-02	0,005	5,08647E-03	-2,997660	1,24505E-02	1,241987				
	1,00520E-02	-2,388006	2,01004E-02	0,01	1,02588E-02	-2,659218	2,47749E-02	1,232556				
	5,13372E-02	-1,669805	1,02586E-01	0,05	5,15155E-02	-1,800643	1,26708E-01	1,235137				
	1,05546E-01	-1,295356	2,10721E-01	0,1	1,05684E-01	-1,371582	2,60098E-01	1,234327				
	7,01217E-01	0,000000	1,38629E+00	0,5	7,06108E-01	0,001447	1,71259E+00	1,235372				
	2,39326E+00	1,295356	4,60518E+00	0,9	2,40492E+00	1,230088	5,66027E+00	1,229111				
	3,15041E+00	1,669805	5,99148E+00	0,95	3,16555E+00	1,566697	7,34661E+00	1,226177				
	4,97744E+00	2,388006	9,21035E+00	0,99	4,99530E+00	2,185930	1,12127E+01	1,217399				
	5,79496E+00	2,657471	1,05965E+01	0,995	5,80295E+00	2,403748	1,28692E+01	1,214473				
	7,76799E+00	3,226960	1,38150E+01	0,999	7,80807E+00	2,855997	1,66598E+01	1,205919				
	8,65111E+00	3,454479	1,52014E+01	0,9995	8,70611E+00	3,041729	1,83790E+01	1,209034				
	1,07830E+01	3,953464	1,84247E+01	0,9999	1,09489E+01	3,435581	2,19775E+01	1,192823				
3x16-21	9,99201E-05	-4,033791	1,99472E-04	0,0001	1,12927E-04	-4,367097	2,28309E-04	1,144567	1,096372	48	63	
	5,00044E-04	-3,509922	9,99722E-04	0,0005	5,06535E-04	-3,740867	1,11520E-03	1,115513				
	1,00053E-03	-3,272871	2,00138E-03	0,001	1,04516E-03	-3,468833	2,26443E-03	1,131436				
	5,01288E-03	-2,684556	1,00247E-02	0,005	4,99884E-03	-2,807396	1,12510E-02	1,122332				
	1,00524E-02	-2,408342	2,01004E-02	0,01	1,00600E-02	-2,506137	2,24383E-02	1,116311				
	5,13518E-02	-1,677927	1,02586E-01	0,05	5,18127E-02	-1,725346	1,14534E-01	1,116463				
	1,05608E-01	-1,299825	2,10721E-01	0,1	1,06114E-01	-1,327085	2,35549E-01	1,117825				
	7,03935E-01	0,000000	1,38629E+00	0,5	7,05099E-01	-0,002829	1,54977E+00	1,117924				
	2,42433E+00	1,299825	4,60518E+00	0,9	2,43441E+00	1,276731	5,12140E+00	1,112097				
	3,20432E+00	1,677927	5,99148E+00	0,95	3,22798E+00	1,640919	6,65052E+00	1,109996				
	5,11034E+00	2,408342	9,21035E+00	0,99	5,16506E+00	2,332648	1,01894E+01	1,106302				
	5,97413E+00	2,684556	1,05965E+01	0,995	6,03800E+00	2,589209	1,16959E+01	1,103745				
	8,08541E+00	3,272871	1,38150E+01	0,999	8,19557E+00	3,129207	1,51560E+01	1,097068				
	9,04220E+00	3,509922	1,52014E+01	0,9995	9,12287E+00	3,334977	1,66704E+01	1,096642				
	1,13796E+01	4,033791	1,84247E+01	0,9999	1,15146E+01	3,812169	1,99473E+01	1,082637				
3x5-5	9,99201E-05	-4,984904	1,99472E-04	0,0001	8,36885E-05	-4,961276	1,69347E-04	1,177889	1,000000	15	15	
	5,00044E-04	-4,140311	9,99722E-04	0,0005	4,66532E-04	-4,163109	9,97908E-04	1,001818				
	1,00053E-03	-3,787427	2,00138E-03	0,001	1,02236E-03	-3,809738	2,00753E-03	0,996936				
	5,01466E-03	-2,976849	1,00247E-02	0,005	5,13110E-03	-2,981930	1,01986E-02	0,982945				
	1,00586E-02	-2,624492	2,01004E-02	0,01	1,01208E-02	-2,624226	2,00350E-02	1,003265				
	5,15135E-02	-1,761309	1,02586E-01	0,05	5,12369E-02	-1,755552	1,03058E-01	0,995427				
	1,06291E-01	-1,345031	2,10721E-01	0,1	1,05610E-01	-1,342623	2,12150E-01	0,993265				
	7,34772E-01	0,000000	1,38629E+00	0,5	7,34956E-01	-0,000754	1,38612E+00	1,000123				
	2,80679E+00	1,345031	4,60518E+00	0,9	2,81044E+00	1,343413	4,58234E+00	1,004983				
	3,88529E+00	1,761309	5,99148E+00	0,95	3,89249E+00	1,761914	5,94820E+00	1,007276				
	6,26249E+00	2,624492	9,21035E+00	0,99	6,94346E+00	2,622087	9,10030E+00	1,012094				
	8,50969E+00	2,976849	1,05965E+01	0,995	8,46008E+00	2,978419	1,04500E+01	1,014020				
	1,29730E+01	3,787427	1,38150E+01	0,999	1,30246E+01	3,793494	1,34899E+01	1,024098				
	1,52977E+01	4,140311	1,52014E+01	0,9995	1,53453E+01	4,174061	1,48686E+01	1,022378				
	2,18570E+01	4,984904	1,84247E+01	0,9999	2,11456E+01	4,942792	1,79786E+01	1,024816				

Docu

	Normal distribution		χ^2	p	Censored normal		χ^2	χ^2 ratio	χ^2 ratio calc	known	total
	F	t			F	t					
5-5,10-10, 15-15	9,99201E-05	-4,253816	1,99472E-04	0,0001	9,61066E-05	-4,252509	2,16112E-04	1,083419	1,000000	30	30
	5,00044E-04	-3,659516	9,99722E-04	0,0005	5,13608E-04	-3,655127	9,96319E-04	0,996596			
	1,00053E-03	-3,396271	2,00138E-03	0,001	1,02626E-03	-3,392013	2,05228E-03	1,025433			
	5,01377E-03	-2,756387	1,00247E-02	0,005	5,09138E-03	-2,756790	1,01703E-02	1,014523			
	1,00542E-02	-2,462020	2,01004E-02	0,01	1,01699E-02	-2,458602	2,02741E-02	1,008643			
	5,13909E-02	-1,699127	1,02586E-01	0,05	5,16352E-02	-1,697078	1,02429E-01	0,998469			
	1,05773E-01	-1,311435	2,10721E-01	0,1	1,05977E-01	-1,308639	2,10748E-01	1,000131			
	7,11250E-01	0,000000	1,38629E+00	0,5	7,11112E-01	0,003785	1,38528E+00	0,999269			
	2,51061E+00	1,311435	4,60518E+00	0,9	2,50538E+00	1,315349	4,59766E+00	0,998368			
	3,35413E+00	1,699127	5,99148E+00	0,95	3,34815E+00	1,703621	5,98251E+00	0,998504			
	5,48812E+00	2,462020	9,21035E+00	0,99	5,49865E+00	2,459101	9,16049E+00	0,994587			
	6,48856E+00	2,756387	1,05965E+01	0,995	6,51702E+00	2,752576	1,05271E+01	0,993451			
	9,01946E+00	3,396271	1,38150E+01	0,999	8,94847E+00	3,400170	1,36814E+01	0,990332			
	1,02063E+01	3,659516	1,52014E+01	0,9995	1,01440E+01	3,662341	1,49253E+01	0,981838			
	1,32131E+01	4,253816	1,84247E+01	0,9999	1,28109E+01	4,277616	1,79778E+01	0,975741			
61-31,41-21, 21-Nov	1,220107	63	1,00005E-04	-3,953818	2,00010E-04		0,0001		1,02530E-04	-4,714394	1,140967
	5,00129E-04	-3,454485	1,00025E-03	0,0005	5,03090E-04	-4,053280	1,14436E-03	1,144070			
	1,00052E-03	-3,226964	2,00100E-03	0,001	1,01519E-03	-3,734841	2,34706E-03	1,172942			
	5,01296E-03	-2,657479	1,00251E-02	0,005	5,08666E-03	-3,005347	1,20659E-02	1,203569			
	1,00520E-02	-2,388011	2,01007E-02	0,01	1,02061E-02	-2,662342	2,45076E-02	1,219244			
	5,13372E-02	-1,669804	1,02587E-01	0,05	5,14018E-02	-1,802515	1,25730E-01	1,225595			
	1,05546E-01	-1,295356	2,10721E-01	0,1	1,06018E-01	-1,375773	2,58551E-01	1,226983			
	7,01217E-01	0,000000	1,38629E+00	0,5	7,03149E-01	-0,001028	1,70463E+00	1,229629			
	2,39325E+00	1,295356	4,60517E+00	0,9	2,39642E+00	1,230685	5,62662E+00	1,221805			
	3,15041E+00	1,669804	5,99146E+00	0,95	3,15459E+00	1,564550	7,29040E+00	1,216798			
	5,97743E+00	2,388011	9,21034E+00	0,99	4,98783E+00	2,181969	1,11180E+01	1,207116			
	6,79499E+00	2,657479	1,05966E+01	0,995	5,80130E+00	2,405222	1,27585E+01	1,204014			
	7,76776E+00	3,226964	1,38155E+01	0,999	7,79348E+00	2,861020	1,65827E+01	1,200296			
	8,65054E+00	3,454485	1,52018E+01	0,9995	8,62823E+00	3,035550	1,83220E+01	1,205250			
	1,07807E+01	3,953818	1,84207E+01	0,9999	1,07430E+01	3,429056	2,18492E+01	1,186124			
4x101-51	1,73397E-03	-3,788068	5,21483E-03	0,0001	1,96705E-03	-4,231153	6,64233E-03	1,273738	1,240173	204	404
	5,08051E-03	-3,339097	1,52790E-02	0,0005	5,10268E-03	-3,627678	1,94585E-02	1,273550			
	8,07953E-03	-3,130863	2,42976E-02	0,001	8,14210E-03	-3,392580	3,12586E-02	1,286488			
	2,38520E-02	-2,600265	7,17218E-02	0,005	2,38567E-02	-2,781175	9,05417E-02	1,262402			
	3,81930E-02	-2,344857	1,14832E-01	0,01	3,77329E-02	-2,493693	1,43872E-01	1,252893			
	1,17093E-01	-1,652394	3,51846E-01	0,05	1,16558E-01	-1,727930	4,36545E-01	1,240726			
	1,94590E-01	-1,285736	5,84374E-01	0,1	1,94343E-01	-1,329171	7,25845E-01	1,242089			
	7,91357E-01	0,000000	2,36597E+00	0,5	7,94208E-01	0,000500	2,93356E+00	1,239897			
	2,11137E+00	1,285736	6,25139E+00	0,9	2,12813E+00	1,252153	7,73430E+00	1,237213			
	2,64975E+00	1,652394	7,81473E+00	0,95	2,67905E+00	1,595910	9,63516E+00	1,232949			
	3,88102E+00	2,344857	1,13449E+01	0,99	3,92905E+00	2,223968	1,40055E+01	1,234526			
	4,40843E+00	2,600265	1,28382E+01	0,995	4,47392E+00	2,456656	1,58596E+01	1,235352			
	5,63414E+00	3,130863	1,62662E+01	0,999	5,74477E+00	2,931311	2,00739E+01	1,234082			
	6,16392E+00	3,339097	1,77300E+01	0,9995	6,29437E+00	3,120429	2,17536E+01	1,226940			
	7,40124E+00	3,788068	2,11075E+01	0,9999	7,65934E+00	3,504696	2,56218E+01	1,213873			

A.6 Key to graphs for composite groups and single sub-groups

Section (File name in disc)	Sub- groups	Detail					
Graph set 1	1	5/5					RAN 3
	1	10/10					RAN 3
	1	21/21					RAN 3
	1	100/100					RAN 3
	1	100/100					RAN 1
	1	200/200					RAN 1
	1	3/5					RAN 3
	1	4/7					RAN 3
	1	5/8					RAN 3
	1	5/9					RAN 3
	1	11/21					RAN 3
	1	15/21					RAN 3
Graph set 2	3	5/5	5/5	5/5			RAN 3
	3	10/10	10/10	10/10			RAN 3
	3	21/21	21/21	21/21			RAN 3
	3	40/40	20/20	10/10			RAN 3
	3	60/60	60/60	60/60			RAN 3
	3	101/101	101/101	101/101			RAN 3
	4	200/200	200/200	200/200	200/200		RAN 3
	4	5/5	5/5	5/5	5/5		RAN 3
Graph set 3	3	21/41	21/41	21/41			RAN 3
	4	26/51	26/51	26/51	26/51		RAN 3
	4	45/51	45/51	45/51	45/51		RAN 3
	4	47/51	47/51	47/51	47/51		RAN 3
	4	50/51	50/51	50/51	50/51		RAN 3
Graph set 4	3	3/5	4/5	5/5			RAN 3
	4	21/21	11/11	8/11	4/7		RAN 3
	4	31/61	31/51	31/41	31/31		RAN 3
	4	11/21	14/21	17/21	21/21		RAN 3
	5	21/21	18/21	16/21	14/21	11/21	RAN 3
Graph set 5	3	6/11	6/11	6/11			RAN 3 See note
	3	6/11	6/11	6/11			RAN 3 See note
	3	11/21	11/21	11/21			RAN 3 See note
	3	11/21	11/21	11/21			RAN 3 See note
	3	31/61	21/41	11/21			RAN 3 See note
	3	31/61	21/41	11/21			RAN 3 See note
	4	31/61	31/51	31/41	31/31		RAN 3 See note
	4	31/61	31/51	31/41	31/31		RAN 3 See note

Section (File name in disc)	Sub- groups	Detail				
Graph set 6	4	5/5	5/5	5/5	5/5	RAN 1
	3	11/11	11/11	11/11		RAN 1
	3	21/21	21/21	21/21	21/21	RAN 1
	3	100/100	100/100	100/100		RAN 1
	4	100/100	100/100	100/100	100/100	RAN 1
	3	200/200	200/200	200/200	200/200	RAN 1
	4	100/100	100/100	100/100		RAN 1BM
	3	100/100	100/100	100/100		RAN 3 BM
Graph set 8	2	41/41	11/21			RAN 3
2	2	11/21	41/41			RAN 3
3	2	14/21	11/21			RAN 3
4	2	11/21	14/21			RAN 3
5	2	21/21	11/21			RAN 3
6	2	11/21	21/21			RAN 3
7	2	11/21	11/21			RAN 3
8	2	8/11	6/11			RAN 3
9	2	3/5	5/5			RAN 3
10	2	4/6	4/6			RAN 3
11	2	5/5	5/5			RAN 3
12	2	26/50	26/50			RAN 3
13	2	16/31	6/11			RAN 3
14	2	41/41	11/21			RAN 3
15	2	16/31	16/16			RAN 3
16	2	10/15	6/9			RAN 3

NOTE Sample pairs in Graph set 6 show the t and χ^2 distributions with the adjustment for censoring applied in separate graphs either to the observed value or to the theoretical function value.

A.7 Graph Sets 7A and 7B

A.7.1 Graph set 7A

Number of groups	Sub-group 1	Sub-group 2	Sub-group 3	Sub-group 4	Slope	Intercept
4	11/21	11/21	11/21	11/210	0	0
4	5/9	5/9	5/9	5/9	0	0
3	6/11	11/21	16/31		0	0
3	11/21	18/21	20/21		0	0

Number of groups	Sub-group 1	Sub-group 2	Sub-group 3	Sub-group 4	Slope	Intercept
3	6/11	6/8	6/6		0	0
3	11/21	11/21	11/20		0	0
3	16/31	16/25	16/19		0	0
3	16/31	16/21	16/16		0	0
3	11/21	16/21	19/20		0	0
3	11/20	16/21	21/21		0	0
3	11/20	16/21	21/21		0	0
3	26/51	26/26	11/21		0	0
3	6/11	11/21	21/21		0	0

A.7.2 Graph set 7B

Number of groups	Sub-group 1	Sub-group 2	Sub-group 3	Sub-group 4	Slope	Intercept
3	5/9	5/9	5/9		0,5	10
3	26/51	26/51	26/51		0,5	10
4	11/21	11/21	11/21	11/21	0,5	10
3	6/11	11/21	16/31		1	1
3	11/21	18/21	20/21		0	0
3	6/11	6/8	6/6		0	0
3	11/21	11/21	11/20		0,5	10
3	16/31	16/25	16/19		1	1
3	16/31	16/23	16/16		0	0
3	11/21	16/21	19/20		0,5	10
3	11/20	16/21	21/21		0,5	10

Number of groups	Sub-group 1	Sub-group 2	Sub-group 3	Sub-group 4	Slope	Intercept
3	11/21	16/21	21/21		1	1
3	26/51	26/26	11/21		1	1
3	6/11	11/21	21/21		0,5	10
3	31/31	31/31	31/31		0,5	103
3	11/11	21/21	31/31		0,5	10

Annex B (informative)

Key to graph features

B.1 General

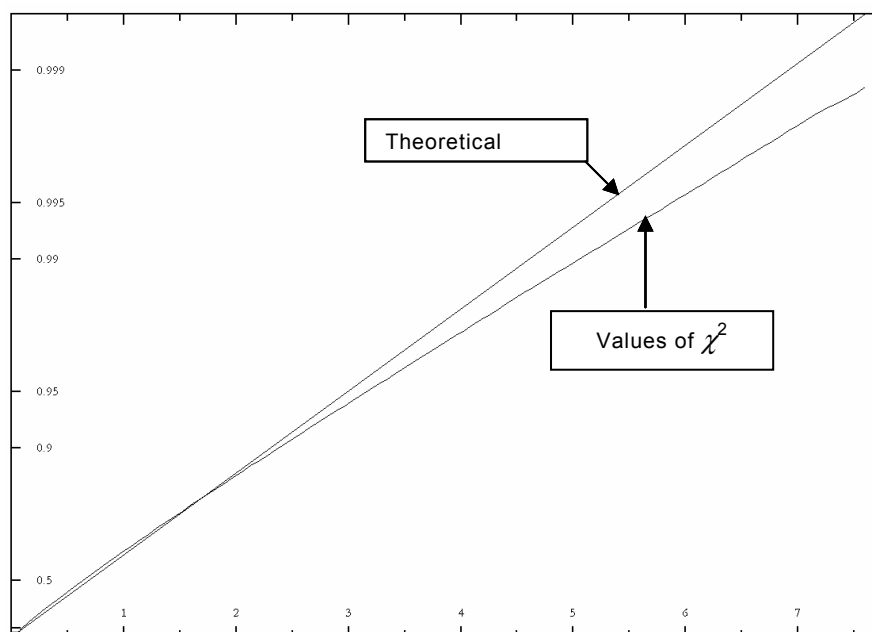
The following clauses state the significance of the various parts of each type of graph.

The scales are calculated so that the graph for the theoretical distribution is a straight line, approximately diagonally placed in the graph.

Please note that the line or curve for the observed values will often coincide with the line for the theoretical distribution. This is especially true of the graphs for complete data.

B.2 Graphs showing mean and variance of single sub-groups

These graphs are contained in the file “Graph Set 1.doc”.



In the graphs for mean, the observed and theoretical lines generally coincide.

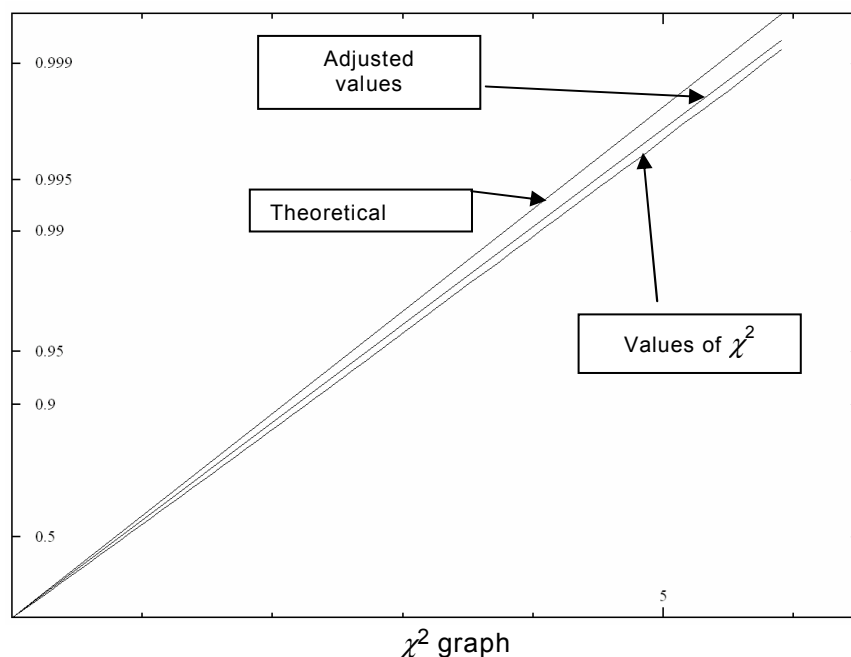
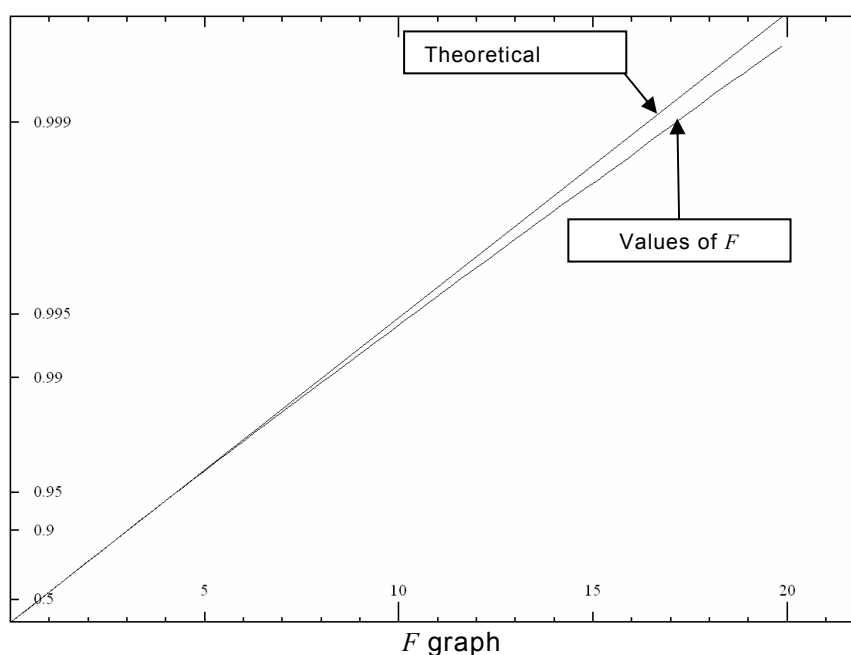
B.3 Graphs showing distributions of F , t and χ^2

B.3.1 General

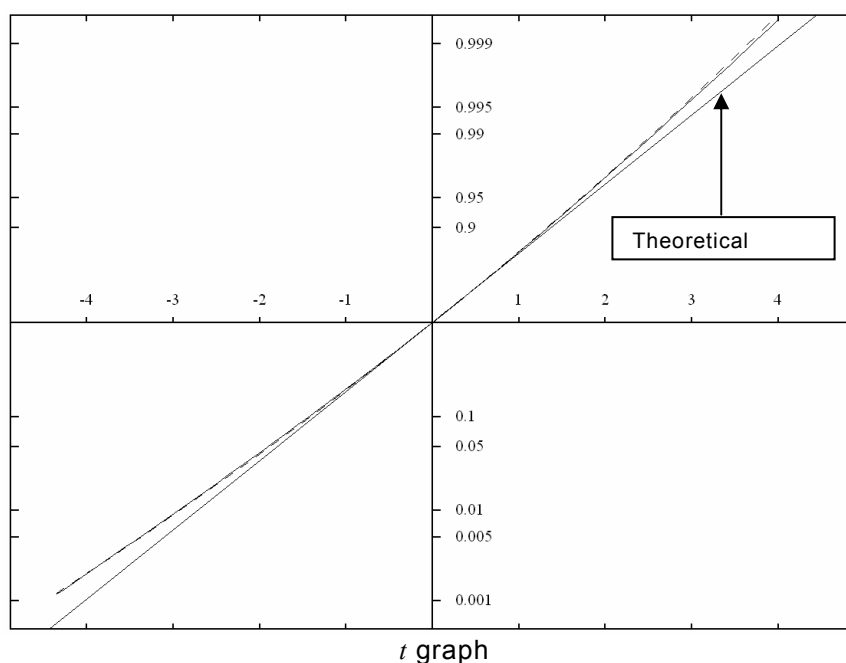
These graphs are in two sections:

- one graph trio for each group tested (file names “Graph Set 2.doc”, Graph Set 3.doc, Graph Set 4.doc and Graph Set 6.doc);
- two graph trios for each group: the first showing the theoretical distribution values adjusted to compare with the observed distribution, the second showing the observed values adjusted for comparison with the theoretical. (Graph Set 5.doc).

B.3.2 Section a)



In the above graph, the adjusted theoretical values are shown for comparison with the observed values.

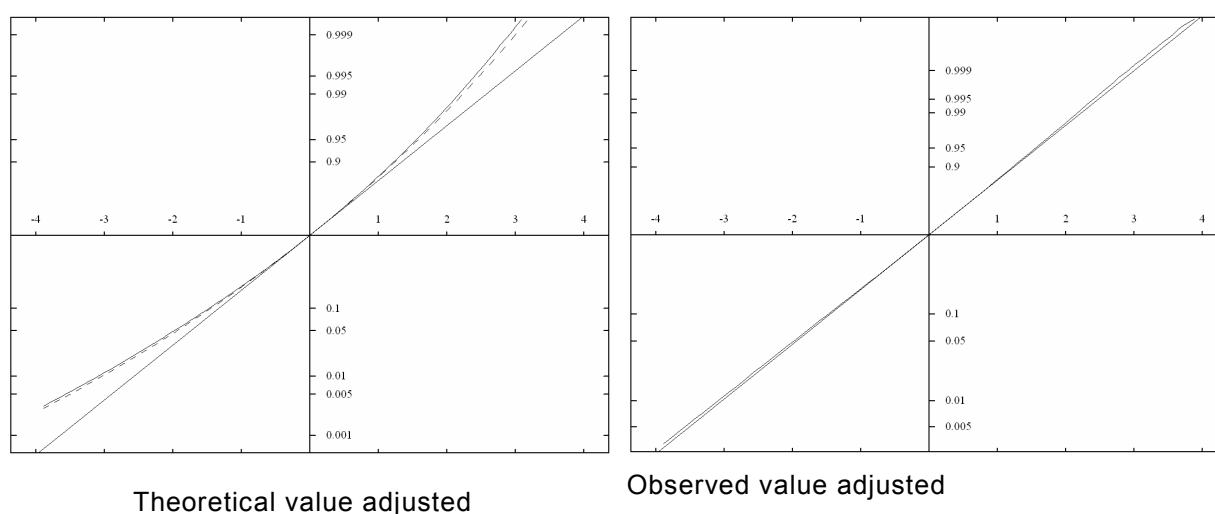


In the t -graph, the observed values and the adjusted theoretical values almost coincide. The adjusted values are shown by the dashed curve.

B.3.3 Section b)

The graphs are in consecutive pairs, the first showing the theoretical distribution values adjusted to compare with the observed distribution, the second showing the observed values adjusted for comparison with the theoretical.

The makeup of the graphs is the same as in Section a), with the above difference between the pairs. The t -graphs are shown as example



B.4 Values of the t -ratio for the difference of two means

Values of the t -ratio for the difference of two means