

DIN ISO 5479



ICS 03.120.30

**Statistische Auswertung von Daten –
Tests auf Abweichung von der Normalverteilung (ISO 5479:1997)**

Statistical interpretation of data –
Tests for departure from the normal distribution (ISO 5479:1997)

Interprétation statistique des données –
Tests pour les écarts à la distribution normale (ISO 5479:1997)

Gesamtumfang 75 Seiten

Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ) im DIN

Die Internationale Norm ISO 5479:1997 „Statistical interpretation of data — Tests for departure from the normal distribution“ ist unverändert in diese Deutsche Norm übernommen worden.

Nationales Vorwort

Die vorliegende Norm hat das ISO/TC 69, *Anwendung statistischer Methoden*, unter Mitwirkung des Normenausschusses Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ) im DIN erarbeitet.

In dieser Norm werden statistische Tests behandelt, mit denen ermittelt wird, ob Messergebnisse einer Stichprobe als aus einer normalverteilten Grundgesamtheit stammend angesehen werden können. Mit den beschriebenen Tests wird die Nullhypothese geprüft, nach der die Messergebnisse der Grundgesamtheit normalverteilt sind. Die dargestellten Methoden unterscheiden sich durch ihre Zielsetzung. Allgemeine Anpassungstests (z. B. Kolmogorow-Smirnow) kommen nicht in Betracht.

Es werden gerichtete Tests behandelt, die entweder allein die Schiefe oder allein die Wölbung bewerten (Begriffe siehe DIN 55350-21). Ferner wird ein mehrfach gerichteter Test beschrieben, der gleichzeitig Schiefe und Wölbung bewertet. Schließlich werden „Omnibus-Tests“ angegeben, die bei beliebiger Abweichung von der Normalverteilung angewendet werden. („Omnibus“ lateinisch: „für alle“, hier im Sinne von „für alle, zumindest für mehrere Parameter der betrachteten Verteilung“.)

Einige Beispiele dienen der Veranschaulichung.

— Leerseite —

Deutsche Übersetzung

Statistische Auswertung von Daten

Tests auf Abweichung von der Normalverteilung

Inhalt

	Seite
Vorwort	6
Einleitung	8
1 Anwendungsbereich	10
2 Normative Verweisungen	10
3 Begriffe und Formelzeichen	12
4 Allgemeines	14
5 Grafische Analyse	16
6 Gerichtete Tests	30
7 Verbundener Test, bei dem $\sqrt{b_1}$ und b_2 gemeinsam angewendet werden (mehrfach gerichteter Test)	38
8 Omnibus-Tests	40
9 Verbundener Test, bei dem mehrere voneinander unabhängige Stichproben angewendet werden	52
10 Statistische Tabellen	56
Anhang A (informativ) Unbenutztes Wahrscheinlichkeitsnetz	72
Anhang B (informativ) Literaturhinweise	74

Statistical interpretation of data
Tests for departure from the normal distribution

Contents

	Page
Foreword.....	7
Introduction	9
1 Scope.....	11
2 Normative references.....	11
3 Definitions and symbols.....	13
4 General	15
5 Graphical method.....	17
6 Directional tests	31
7 Joint test using $\sqrt{b_1}$ and b_2 (multidirectional test).....	39
8 Omnibus-tests	41
9 Joint test using several independent samples.....	53
10 Statistical tables	57
Annex A (informative) Blank normal probability graph paper.....	73
Annex B (informative) Bibliography.....	75

Vorwort

Die ISO (Internationale Organisation für Normung) ist die weltweite Vereinigung nationaler Normungsinstitute (ISO-Mitglieds Körperschaften). Die Erarbeitung Internationaler Normen obliegt den Technischen Komitees der ISO. Jede Mitglieds Körperschaft, die sich für ein Thema interessiert, für das ein Technisches Komitee eingesetzt wurde, ist berechtigt, in diesem Komitee mitzuarbeiten. Internationale (staatliche und nichtstaatliche) Organisationen, die mit der ISO in Verbindung stehen, sind an den Arbeiten ebenfalls beteiligt. Die ISO arbeitet in allen Angelegenheiten der elektrotechnischen Normung eng mit der Internationalen Elektrotechnischen Kommission (IEC) zusammen.

Die von den Technischen Komitees verabschiedeten internationalen Norm-Entwürfe werden den Mitglieds Körperschaften zur Abstimmung vorgelegt. Die Veröffentlichung als Internationale Norm erfordert Zustimmung von mindestens 75 % der abstimmenden Mitglieds Körperschaften.

Die Internationale Norm ISO 5479 wurde durch das Technische Komitee ISO/TC 69, *Anwendung statistischer Methoden*, Unterkomitee SC 6, *Messmethoden und Messergebnisse*, erarbeitet.

Die Anhänge A und B dienen nur zur Information.

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

International Standard ISO 5479 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

Annexes A and B of this International Standard are for information only.

Einleitung

Viele der statistischen Methoden, die in Internationalen Normen empfohlen werden, wie zum Beispiel in ISO 2854 (siehe [1]), beruhen auf der Annahme, dass die Zufallsvariable(n), auf die diese Methoden angewendet werden, voneinander unabhängig und entsprechend einer Normalverteilung verteilt sind, wobei einer oder beide Parameter der Normalverteilung unbekannt sind.

Damit stellt sich die folgende Frage: Liegt die Verteilung, die durch die Stichprobe repräsentiert wird, hinreichend nahe an der Normalverteilung, damit die in diesen Internationalen Normen vorgeschlagenen Methoden vertrauenswürdig angewendet werden können?

Diese Frage kann nicht einfach mit „Ja“ oder „Nein“ beantwortet werden, und sie (die Antwort) gilt nicht für alle (Anwendungs-)Fälle. Deshalb wurde eine große Anzahl von „Tests auf Normalität“ entwickelt, von denen jeder auf ein besonderes Merkmal der betrachteten Verteilung mehr oder weniger empfindlich anspricht, zum Beispiel auf die Asymmetrie oder die Wölbung.

Im Allgemeinen ist der angewendete Test entwickelt worden, um mit einer vorgegebenen *a priori*-Wahrscheinlichkeit die „Hypothese der Normalität“ zu verwerfen, auch wenn sie zutrifft (Fehler 1. Art). Andererseits kann die Wahrscheinlichkeit, dass diese Hypothese nicht verworfen wird, obwohl sie nicht zutrifft (Fehler 2. Art), nur dann ermittelt werden, wenn die Alternativhypothese (d. h. die der „Hypothese der Normalität“ entgegengesetzte Hypothese) genau definiert werden kann. Dies ist im Allgemeinen nicht möglich, und es erfordert darüber hinaus rechnerischen Aufwand. Für einen betrachteten Test ist diese Wahrscheinlichkeit (des Fehlers 2. Art) besonders groß, wenn der Stichprobenumfang klein ist.

Introduction

Many of the statistical methods recommended in International Standards, such as those described in ISO 2854^[1], are based on the assumption that the random variable(s) to which these methods apply are independently distributed according to a normal distribution with one or both of its parameters unknown.

The following question therefore arises. Is the distribution that is represented by the sample sufficiently close to the normal distribution that the methods provided by these International Standards can be used reliably?

There is no simple yes or no answer to this question which is valid in all cases. For this reason a large number of "tests of normality" have been developed, each of which is more or less sensitive to a particular feature of the distribution under consideration; e.g. asymmetry or kurtosis.

Generally the test used is designed to correspond to a predetermined *a priori* risk that the hypothesis of normality is rejected even if it is true (error of the first kind). On the other hand, the probability that this hypothesis is not rejected when it is not true (error of the second kind) cannot be determined unless the alternative hypothesis (i.e. that which is opposed to the hypothesis of normality) can be precisely defined. This is not possible in general and, furthermore, it requires computational effort. For a distinct test, this risk is particularly large if the sample size is small.

1 Anwendungsbereich

1.1 Diese Internationale Norm gibt eine Anleitung für Methoden und Tests, die bei Anwendung dafür ausschlaggebend sind, ob die Hypothese Normalverteilung zu verwerfen ist oder nicht, und zwar unter der Voraussetzung, dass die Messwerte voneinander unabhängig sind.

1.2 Falls Zweifel bestehen, ob die Messwerte normalverteilt sind, kann die Anwendung eines Tests auf Abweichung von der Normalverteilung zweckmäßig oder erforderlich sein. Bei robusten Methoden jedoch (also bei solchen, bei denen die Ergebnisse nur wenig beeinflusst werden, wenn die tatsächliche Wahrscheinlichkeitsverteilung der Messwerte keine Normalverteilung ist) ist ein Test auf Abweichung von der Normalverteilung weniger nützlich. Das ist beispielsweise der Fall, wenn der Mittelwert einer einzelnen Zufallsstichprobe von Messwerten mit einem gegebenen theoretischen Wert mittels eines t-Tests (Student-Tests) verglichen werden soll.

1.3 Es ist nicht unbedingt erforderlich, einen derartigen Test jedes Mal anzuwenden, wenn man auf statistische Methoden zurückgreift, die auf der Annahme der Normalverteilung beruhen. Es kann durchaus zweifelsfrei feststehen, dass die Messwerte normalverteilt sind, entweder weil theoretische (zum Beispiel physikalische) Gründe vorliegen, die diese Hypothese rechtfertigen, oder weil diese Hypothese aufgrund von in der Vergangenheit gewonnenen Informationen als annehmbar gilt.

1.4 Die in dieser Internationalen Norm ausgewählten Tests auf Abweichung von der Normalverteilung sind in erster Linie gedacht für einzeln ermittelte Messwerte, nicht für klassierte Messwerte. Die Tests sind ungeeignet für zensurierte Messwerte.

1.5 Die in dieser Internationalen Norm ausgewählten Tests auf Abweichung von der Normalverteilung sind anwendbar entweder auf die Messwerte selbst oder auf Funktionen der Messwerte, zum Beispiel auf deren Logarithmen oder deren Quadratwurzeln.

1.6 Tests auf Abweichung von der Normalverteilung sind fast wirkungslos bei einem Stichprobenumfang von weniger als acht. Folglich ist diese Internationale Norm eingeschränkt auf Stichprobenumfänge von acht oder mehr.

2 Normative Verweisungen

Die folgenden normativen Dokumente enthalten Festlegungen, die durch Verweisung in diesem Text Bestandteil dieser Internationalen Norm sind. Bei datierten Verweisungen gelten spätere Änderungen oder Überarbeitungen dieser Publikationen nicht. Anwender dieser Internationalen Norm werden jedoch gebeten, die Möglichkeit zu prüfen, die jeweils neuesten Ausgaben der nachfolgend angegebenen normativen Dokumente anzuwenden. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen normativen Dokuments. Mitglieder von IEC und ISO führen Verzeichnisse der gültigen Internationalen Normen.

ISO 3534-1:1993, *Statistik — Begriffe und Formelzeichen — Teil 1: Wahrscheinlichkeit und allgemeine statistische Begriffe*.

1 Scope

1.1 This International Standard gives guidance on methods and tests for use in deciding whether or not the hypothesis of a normal distribution should be rejected, assuming that the observations are independent.

1.2 Whenever there are doubts as to whether the observations are normally distributed, the use of a test for departure from the normal distribution may be useful or even necessary. In the case of robust methods, however (i.e. where the results are only altered very slightly when the real probability distribution of the observations is not a normal distribution), a test for departure from the normal distribution is not very helpful. This is the case, for example, when the mean of a single random sample of observations is to be checked against a given theoretical value using a *t*-test.

1.3 It is not strictly necessary to use such a test whenever one refers to statistical methods based on the hypothesis of normality. It is possible that there is no doubt at all as to the normal distribution of the observations, whether theoretical (e.g. physical) reasons are present which confirm the hypothesis or because this hypothesis is deemed to be acceptable according to prior information.

1.4 The tests for departure from the normal distribution selected in this International Standard are primarily intended for complete data, not grouped data. They are unsuitable for censored data.

1.5 The tests for departure from the normal distribution selected in this International Standard may be applied either to observed values or to functions of them, such as the logarithm or the square root.

1.6 Tests for departure from the normal distribution are very ineffective for samples of size less than eight. Accordingly, this International Standard is restricted to samples of eight or more.

2 Normative reference

The following standard contains provisions which, through reference in this text, constitute provisions of this International Standard. At the time of publication, the edition indicated was valid. All standards are subject to revision, and parties to agreements based on this International Standard are encouraged to investigate the possibility of applying the most recent edition of the standard indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 3534-1:1993, *Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms*.

3 Begriffe und Formelzeichen

3.1 Begriffe

Für die Anwendung dieser Norm gelten die in ISO 3534-1 angegebenen Begriffe.

3.2 Formelzeichen

a_k	Koeffizient für den Shapiro-Wilk-Test
A	Hilfsgröße für den Epps-Pulley-Test
b_2	empirische Wölbung
$\sqrt{b_1}$	empirische Schiefe
B	Hilfsgröße für den Epps-Pulley-Test
E	Erwartungswert
G_j	Hilfsgröße für den verbundenen Test, bei dem mehrere voneinander unabhängige Stichproben angewendet werden
h	Anzahl der aufeinander folgenden Proben
H_0	Nullhypothese
H_1	Alternativhypothese
k	Nummer des Messwertes x innerhalb der in nicht-absteigender Folge geordneten Stichproben-Messwerte
m_j	zentrales Moment der Ordnung j der Stichprobe
n	Stichprobenumfang
p	zum p -Quantil einer Verteilung gehörende Wahrscheinlichkeit
P	Wahrscheinlichkeit
P_k	zu $X_{(k)}$ gehörende Wahrscheinlichkeit
S	Hilfsgröße für den Shapiro-Wilk-Test
T	Prüfgröße
T_{EP}	Prüfgröße für den Epps-Pulley-Test
u_p	p -Quantil der standardisierten Normalverteilung
v_j	Hilfsgröße für den verbundenen Test, bei dem mehrere voneinander unabhängige Stichproben angewendet werden
W	Prüfgröße für den Shapiro-Wilk-Test
W_j	Hilfsgröße für den verbundenen Test, bei dem mehrere voneinander unabhängige Stichproben angewendet werden
x	Messwert von X
X	Zufallsvariable
$x_{(j)}$	Messwert j innerhalb der in nicht-absteigender Folge geordneten Stichproben-Messwerte
$x_{(k)}$	Messwert k innerhalb der in nicht-absteigender Folge geordneten Stichproben-Messwerte
$\bar{x}$	arithmetischer Mittelwert
α	Signifikanzniveau

3 Definitions and symbols

3.1 Definitions

For the purposes of this International Standard, the definitions given in ISO 3534-1 apply.

3.2 Symbols

a_k	coefficient of the Shapiro-Wilk test
A	auxiliary quantity for the Epps-Pulley test
b_2	empirical kurtosis
$\sqrt{b_1}$	empirical skewness
B	auxiliary quantity for the Epps-Pulley test
E	expectation
G_j	auxiliary quantity for the joint test using several independent samples
h	number of consecutive samples
H_0	null hypothesis
H_1	alternative hypothesis
k	within the sample, arranged in non-decreasing order, the number of the observed value x
m_j	central moment of order j of the sample
n	sample size
p	probability associated with the p -quantile of a distribution
P	probability
P_k	probability associated with $X_{(k)}$
S	auxiliary quantity for the Shapiro-Wilk test
T	test statistic
T_{EP}	test statistic of the Epps-Pulley test
u_p	p -quantile of the standardized normal distribution
v_j	auxiliary quantity for the joint test using several independent samples
W	test statistic of the Shapiro-Wilk test
W_j	auxiliary quantity for the joint test using several independent samples
x	value of X
X	random variable
$x_{(j)}$	j^{th} value in the sample, arranged in non-decreasing order
$x_{(k)}$	k^{th} value in the sample, arranged in non-decreasing order
$\bar{x}$	arithmetic average
α	significance level

β	Wahrscheinlichkeit des Fehlers 2. Art
β_2	Wölbung der Grundgesamtheit
$\beta_2 - 3$	Exzess der Grundgesamtheit
$\sqrt{\beta_1}$	Schiefe der Grundgesamtheit
γ	Hilfsgröße für den verbundenen Test, bei dem mehrere voneinander unabhängige Stichproben angewendet werden
$\gamma_{(n)}$	Koeffizient für den verbundenen Test, bei dem mehrere voneinander unabhängige Stichproben angewendet werden
δ	Hilfsgröße für den verbundenen Test, bei dem mehrere voneinander unabhängige Stichproben angewendet werden
$\delta_{(n)}$	Koeffizient für den verbundenen Test, bei dem mehrere voneinander unabhängige Stichproben angewendet werden
ε	Hilfsgröße für den verbundenen Test, bei dem mehrere voneinander unabhängige Stichproben angewendet werden
$\varepsilon_{(n)}$	Koeffizient für den verbundenen Test, bei dem mehrere voneinander unabhängige Stichproben angewendet werden
μ	Erwartungswert (der Grundgesamtheit)
μ_2	Varianz der Grundgesamtheit ^{N1)}
μ_3	zentrales Moment der Ordnung 3 der Grundgesamtheit
μ_4	zentrales Moment der Ordnung 4 der Grundgesamtheit
σ	Standardabweichung der Grundgesamtheit ($= \sqrt{\mu_2}$)

4 Allgemeines

4.1 Es gibt verschiedene Kategorien von Tests auf Abweichung von der Normalverteilung. In dieser Internationalen Norm werden grafische Analysen, Momenten-Tests, Regressions-Tests und Tests mittels der charakteristischen Funktion betrachtet. Chi-Quadrat-Tests sind nur für klassierte Messwerte geeignet; diese Tests werden in dieser Internationalen Norm jedoch nicht betrachtet, weil durch die Klassierung ein Verlust an Informationen entsteht.

4.2 Falls keine zusätzlichen Informationen über die Stichprobe vorliegen, wird empfohlen, zuerst die Messwerte im Wahrscheinlichkeitsnetz grafisch darzustellen, d. h. die Verteilungsfunktion der Messwerte im Wahrscheinlichkeitsnetz (mit linear geteilter Abszisse) aufzuzeichnen. Das Wahrscheinlichkeitsnetz besteht aus einem Koordinatensystem, in dem die Verteilungsfunktion der Normalverteilung durch eine Gerade wiedergegeben ist.

Diese in Abschnitt 5 beschriebene Analyse ermöglicht es, unmittelbar zu „sehen“, ob die beobachtete Verteilung (der Messwerte) in der Nähe der Normalverteilung liegt oder nicht. Mit dieser zusätzlichen Information kann entschieden werden, ob ein gerichteter Test durchzuführen ist, oder ob ein Regressions-Test oder ein Test mittels der charakteristischen Funktion, oder ob gar kein Test durchzuführen ist. Obwohl diese grafische Darstellung nicht als Test im strengen Sinne angesehen werden kann, ist die dadurch erhaltene zusätzliche Übersichtsinformation eine wichtige Ergänzung für jeden Test auf Abweichung von der Normalverteilung. Falls die Nullhypothese verworfen wird, ist es oft möglich, aus dieser grafischen Darstellung Hinweise für das geeignet erscheinende weitere Vorgehen zu entnehmen.

N1) Nationale Fußnote: Die Varianz ist das zentrale Moment der Ordnung 2.

β	probability of an error of the second kind
β_2	kurtosis of the population
β_2-3	excess of the population
$\sqrt{\beta_1}$	skewness of the population
γ	auxiliary quantity for the joint test using several independent samples
$\gamma_{(n)}$	coefficient of the joint test using several independent samples
δ	auxiliary quantity for the joint test using several independent samples
$\delta_{(n)}$	coefficient of the joint test using several independent samples
ε	auxiliary quantity for the joint test using several independent samples
$\varepsilon_{(n)}$	coefficient of the joint test using several independent samples
μ	expectation
μ_2	variance of the population
μ_3	central moment of the third order of the population
μ_4	central moment of the fourth order of the population
σ	standard deviation of the population ($=\sqrt{\mu_2}$)

4 General

4.1 There are several categories of tests for departure from normality. In this International Standard, graphical methods, moment tests, regression tests and characteristic function tests are considered. Chi-squared tests are appropriate for grouped data only but, because grouping results in a loss of information, they are not considered in this International Standard.

4.2 If no additional information about the sample is available, it is recommended first to do a normal probability plot; i.e. to plot the cumulative distribution function of the observed values on normal probability graph paper consisting of a system of coordinate axes where the cumulative distribution function of the normal distribution is represented by a straight line.

This method, which is described in clause 5, allows one to “see” immediately whether the distribution observed is close to the normal distribution or not. With this additional information it can be decided whether to carry out a directional test, or to carry out either a regression test or a characteristic function test, or no test at all. In addition, although such a graphical representation cannot be considered as a rigorous test, the summary information that it provides is an essential supplement to any test for departure from the normal distribution. In the case of rejection of the null hypothesis it is often possible to envisage by this means the type of alternative that might be applicable.

4.3 A test for departure from the normal distribution is a test of the null hypothesis that the sample consists of n independent observations coming from one and the same normal distribution. It consists of the calculation of a function T of the observations, which is called the test statistic. The null hypothesis of a normal distribution is then not rejected or rejected depending on whether or not the value of T lies within a set of values near to the expected value that corresponds to the normal distribution.

4.3 Ein Test auf Abweichung von der Normalverteilung ist ein Test der Nullhypothese, dass die Stichprobe aus n voneinander unabhängigen Messwerten besteht, die aus ein und derselben Normalverteilung stammen. Dieser Test besteht in der Berechnung einer Funktion T aus den Messwerten, die „Prüfgröße“ genannt wird. Die Nullhypothese Normalverteilung wird nicht verworfen oder verworfen, je nachdem, ob der Wert der Prüfgröße T innerhalb oder nicht innerhalb eines Wertebereiches in der Umgebung des Erwartungswertes liegt, welcher der Normalverteilung entspricht.

4.4 Der **kritische Bereich** des Tests umfasst die Menge aller Werte T , die zum Verwerfen der Nullhypothese führt. Das **Signifikanzniveau** des Tests ist die Wahrscheinlichkeit P , mit der man einen Wert T innerhalb des kritischen Bereiches erhält, wenn die Nullhypothese nicht verworfen wird. Dieses Signifikanzniveau entspricht der Wahrscheinlichkeit, die Nullhypothese irrtümlicherweise zu verwerfen (Fehler 1. Art).

Die Grenze des kritischen Bereiches ist der kritische Wert der Prüfgröße, oder die Grenzen des kritischen Bereiches beim zweiseitigen Test sind die kritischen Werte der Prüfgröße.

4.5 Die **Schärfe** des Tests ist die Wahrscheinlichkeit, mit der die Nullhypothese verworfen wird, wenn sie nicht zutrifft. Eine große Schärfe entspricht einer niedrigen Wahrscheinlichkeit, die Nullhypothese irrtümlich nicht zu verwerfen (Fehler 2. Art).

Es ist hervorzuheben, dass die Schärfe eines Tests (d. h., unter gegebenen Bedingungen die Wahrscheinlichkeit, dass die Nullhypothese Normalverteilung verworfen werden wird, wenn sie nicht zutrifft) mit der Anzahl von Messwerten zunimmt. Zum Beispiel kann eine Abweichung von der Normalverteilung offenbar werden, wenn ein Test auf Abweichung von der Normalverteilung auf eine große Stichprobe angewendet wird, wogegen mit demselben Test keine Abweichung von der Normalverteilung entdeckt wird, wenn weniger Messwerte vorliegen.

4.6 Bei Tests auf Abweichung von der Normalverteilung sind zwei Kategorien von Tests zu unterscheiden. Wenn die Art der Abweichung von der Normalverteilung in der Alternativhypothese festgelegt ist, handelt es sich um einen **gerichteten Test**. Wenn jedoch die Art der Abweichung von der Normalverteilung in der Alternativhypothese nicht festgelegt ist, handelt es sich um einen **Omnibus-Test**.

Bei einem gerichteten Test wird der kritische Bereich so ermittelt, dass die Schärfe des Tests ihren Maximalwert annimmt. Bei einem Omnibus-Test ist es erforderlich, den kritischen Bereich so aufzuteilen, dass der kritische Bereich aus jenen Werten der Prüfgröße besteht, die vom Erwartungswert weit entfernt liegen.

Falls Annahmen über die Art der Abweichung von der Normalverteilung vorliegen, d. h., wenn eine Verteilung in Betracht gezogen wird, deren Asymmetrie oder Wölbung sich von der einer Normalverteilung unterscheidet, sollte ein gerichteter Test angewendet werden, weil dessen Schärfe größer ist als die Schärfe eines Omnibus-Tests.

4.7 Es ist zu beachten, dass ein gerichteter Test von Natur aus ein einseitiger Test ist. Beispielsweise richtet sich der Test bei einer Asymmetrie entweder auf eine positive oder auf eine negative Asymmetrie. Werden jedoch mehrere Alternativen gemeinsam betrachtet, ist der Test ein mehrfach gerichteter Test. Dieses ist insbesondere der Fall, wenn eine von Null abweichende Asymmetrie und eine Wölbung, die von derjenigen bei einer Normalverteilung abweicht, gemeinsam betrachtet werden.

4.8 Die Tabellen 8 bis 14 und Bild 9 erlauben es, die Tests für die gebräuchlichsten Signifikanzniveaus $\alpha = 0,05$ und $\alpha = 0,01$ durchzuführen. Das Signifikanzniveau α ist **vor** der Durchführung des Tests festzulegen. Es ist zu beachten, dass sich aus einem Test das Verwerfen der Nullhypothese auf dem Signifikanzniveau 0,05 und das Nicht-Verwerfen derselben Hypothese auf dem Signifikanzniveau 0,01 ergeben kann.

4.9 Während der Berechnung von Prüfgrößen ist es erforderlich, mindestens sechs signifikante Stellen anzuwenden. Zwischensummen, Zwischenergebnisse und Hilfsgrößen dürfen nicht auf weniger als sechs signifikante Stellen gerundet werden.

5 Grafische Analyse

5.1 Die empirische Verteilungsfunktion der Messwerte wird im Wahrscheinlichkeitsnetz grafisch dargestellt. In diesem Netz ist eine der Achsen (in dieser Internationalen Norm die vertikale Achse) nicht-linear gemäß der

4.4 The **critical region** of the test is the set of values of T that leads to the rejection of the null hypothesis. The **significance level** of the test is the probability P of obtaining a value of T within the critical region when the null hypothesis is correct. This level gives the probability of erroneously rejecting the null hypothesis (error of the first kind).

The boundary of the critical region is (or, in the case of a two-sided test, the boundaries of the critical region are) the critical value(s) of the test statistic.

4.5 The **power** of the test is the probability of rejecting the null hypothesis when it is incorrect. A high power corresponds to a low probability of not rejecting the null hypothesis erroneously (error of the second kind).

It should be emphasized that the power of a test (i.e. for a given situation, the probability that the null hypothesis of a normal distribution will be rejected if it is wrong) increases as the number of observations increases. For example, a departure from the normal distribution which would become apparent when using a test for departure from the normal distribution on a large sample might not be detected by the same test if there were fewer observations.

4.6 A distinction is made between two categories of tests for departure from the normal distribution. When the form of departure from the normal distribution is specified in the alternative hypothesis, then the test is a **directional test**. However, when the form of departure from the normal distribution is not specified in the alternative hypothesis, the test is an **omnibus test**.

In a directional test, the critical region is determined in such a way that the power of the test reaches its maximum value. In an omnibus test, it is necessary to divide the critical region in such a way that the critical region consists of those values of the test statistic which lie far away from the expected value.

If assumptions are present about the type of departure from the normal distribution, i.e. when a distribution is envisaged whose asymmetry or kurtosis is different from that of the normal distribution, a directional test should be applied, because its power is greater than the power of an omnibus test.

4.7 Note that a directional test is essentially one-sided. In the case of asymmetry, for example, it centres either on positive asymmetry or on negative asymmetry. However, when several alternatives are considered jointly, the test is multidirectional. This is the case particularly when a non-null asymmetry and a kurtosis different from that of the normal distribution are considered together.

4.8 Tables 8 to 14 and figure 9 allow the tests to be performed for the most usual levels of α ; i.e. $\alpha = 0,05$ and $\alpha = 0,01$. The level of significance has to be stipulated before the test is performed. Note that a test may result in the rejection of the null hypothesis at the 0,05 level and the non-rejection of this same hypothesis at the 0,01 level.

4.9 During computation of test statistics, it is necessary to use at least six significant digits. Subtotals, intermediate results and auxiliary quantities shall not be rounded to less than six significant digits.

5 Graphical method

5.1 The cumulative distribution function of the observed values is plotted on normal probability graph paper. On this paper, one of the axes (in this International Standard it is the vertical axis) is non-linearly scaled according to the area under the standardized normal distribution function and is marked with the corresponding values of the cumulative relative frequency. The other axis is linearly scaled for the ordered values of X . The cumulative distribution function of the variable X then approximates to a straight line.

Sometimes these two axes are interchanged with each other. Furthermore, if a normalizing transformation of the variable X is made, the linear scale may be replaced by a logarithmic, quadratic, reciprocal or other scale.

Figure 1 gives an example of normal probability graph paper. On the vertical axis the values of the cumulative relative frequency are given as percentages, while the horizontal axis has an arbitrary linear scale.

Fläche unter der standardisierten Normalverteilung geteilt und mit den entsprechenden Werten der kumulierten relativen Häufigkeit bezeichnet. Die andere Achse ist linear geteilt für die geordneten Werte von X . Die empirische Verteilungsfunktion der Variablen X ist näherungsweise eine Gerade.

Gelegentlich sind diese beiden Achsen miteinander vertauscht. Außerdem kann die lineare Teilung durch eine logarithmische, quadratische, reziproke oder andere Teilung ersetzt werden, wenn die Variable X normalisierend transformiert wird.

Bild 1 zeigt ein Beispiel für das Wahrscheinlichkeitsnetz. Auf der vertikalen Achse sind die Werte der kumulierten relativen Häufigkeit in Prozent angegeben, die horizontale Achse ist in willkürlichen Einheiten linear geteilt.

Ein unbenutztes Blatt des Wahrscheinlichkeitsnetzes ist in Anhang A beigelegt.

Falls die grafische Darstellung eine Folge von Punkten ergibt, die um eine annähernd gerade Linie streuen, gibt das einen rohen Hinweis auf die Annahme, dass die Stichprobe als aus einer Normalverteilung stammend angesehen werden kann.

Falls die Folge von Punkten jedoch von einer geraden Linie systematisch abweicht, gibt die grafische Darstellung oft einen Hinweis auf den Verteilungstyp, der in Betracht zu ziehen ist.

Die Bedeutung dieses Ansatzes liegt in der leicht zugänglichen visuellen Information über die Art der Abweichung von der Normalverteilung.

Falls die grafische Darstellung darauf hinweist, dass die Messwerte aus einer verformten Verteilung stammen (beispielsweise wenn die grafische Darstellung wie in den Bildern 5 oder 6 verläuft), kann eine Transformation der Messwerte zu einer Normalverteilung führen.

Falls die grafische Darstellung darauf hinweist, dass die Messwerte nicht aus einer einfachen homogenen Verteilung stammen, sondern eher aus einer Überlagerung von zwei oder mehreren homogenen Unterverteilungen (beispielsweise wenn die grafische Darstellung wie in Bild 7 verläuft), ist es empfehlenswert, die Unterverteilungen zu identifizieren und die Untersuchung für jede Unterverteilung getrennt weiterzuführen.

Es sollte stets beachtet werden, dass eine derartige grafische Darstellung keinesfalls ein Test auf Abweichung von der Normalverteilung im strengen Sinne ist. Bei Stichproben mit geringem Umfang können deutlich gekrümmte Kurven auch bei Normalverteilungen entstehen, während bei Stichproben mit großem Umfang bereits schwach gekrümmte Kurven auf nicht-normale Verteilungen hinweisen können.

5.2 Bei dem grafischen Verfahren werden die Messwerte $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$ in nicht-abnehmender Folge geordnet. Anschließend wird

$$P_k = (k - 3/8) / (n + 1/4)$$

gegen $x_{(k)}$ im Wahrscheinlichkeitsnetz aufgetragen.

ANMERKUNG Häufig benutzte Alternativen zu Gleichung 1 sind

$$P_k = (k - 1/2) / n$$

und

$$P_k = k / (n + 1)$$

Diese Alternativen sind weniger günstigen Näherungen an die Normalverteilung der erwarteten Ranggröße $F[E(X_{(k)})]$. Ihre Anwendung wird nicht empfohlen.

5.3 Bild 2 zeigt ein Beispiel zur Anwendung des Wahrscheinlichkeitsnetzes.

A sheet of blank normal probability graph paper is provided in annex A.

If a plot on this paper gives a set of points that appears to be scattered around a straight line, this provides crude support for the assumption that the sample can reasonably be regarded as having come from a normal distribution.

However, if there is a systematic departure from the straight line, the plot often suggests the type of distribution to be taken into consideration.

The importance of this approach is that it easily provides visual information on the type of departure from the normal distribution.

If the graph indicates that the data come from a shaped distribution (e.g. if the graph of the cumulative distribution function is as shown in figure 5 or 6), a transformation of the data might result in a normal distribution.

If the graph indicates that the data do not come from a simple homogeneous distribution, but rather from a mixture of two or more homogeneous subpopulations (e.g. if the graph of the cumulative distribution function is as shown in figure 7), it is recommended that the subpopulations be identified and the analysis on each subpopulation be continued separately.

It should be kept in mind that such a plot is in no way a test for departure from the normal distribution in the strict sense. In the case of small samples, pronounced curves may occur for normal distributions, whilst for large samples slight curves may indicate non-normal distributions.

5.2 The graphical procedure consists of arranging the observed values $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$ in non-decreasing order, and then plotting

$$P_k = (k - 3/8)/(n + 1/4) \quad \dots (1)$$

against $x_{(k)}$ on normal probability graph paper.

NOTE 1 Commonly used alternatives to equation (1) are

$$P_k = (k - 1/2)/n$$

and

$$P_k = k/(n + 1)$$

These are poorer approximations to the normal distribution function of the expected order statistics, $F[E(X_{(k)})]$, and their use is not recommended.

5.3 An example of how normal probability graph paper is used is shown in figure 2.

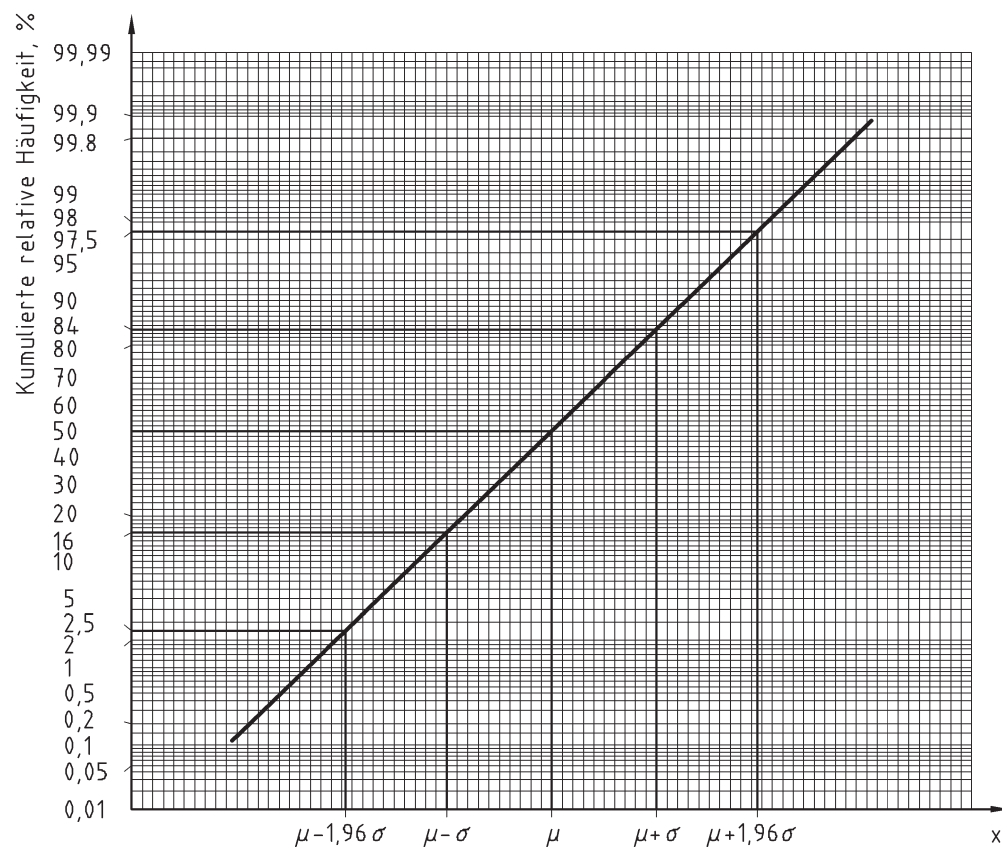


Bild 1 — Mit Hinweisen versehenes Wahrscheinlichkeitsnetz

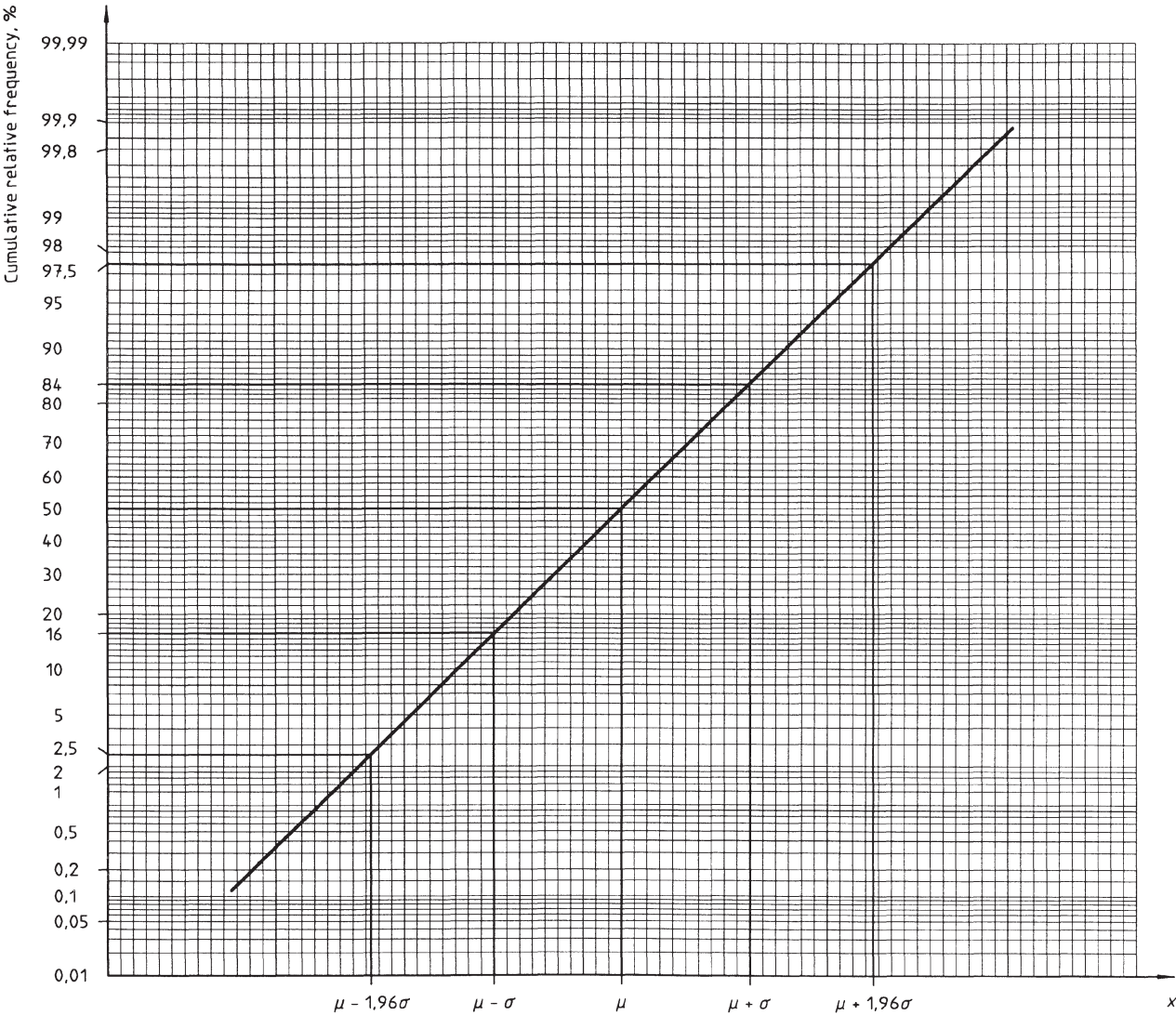


Figure 1 — Annotated normal probability graph paper

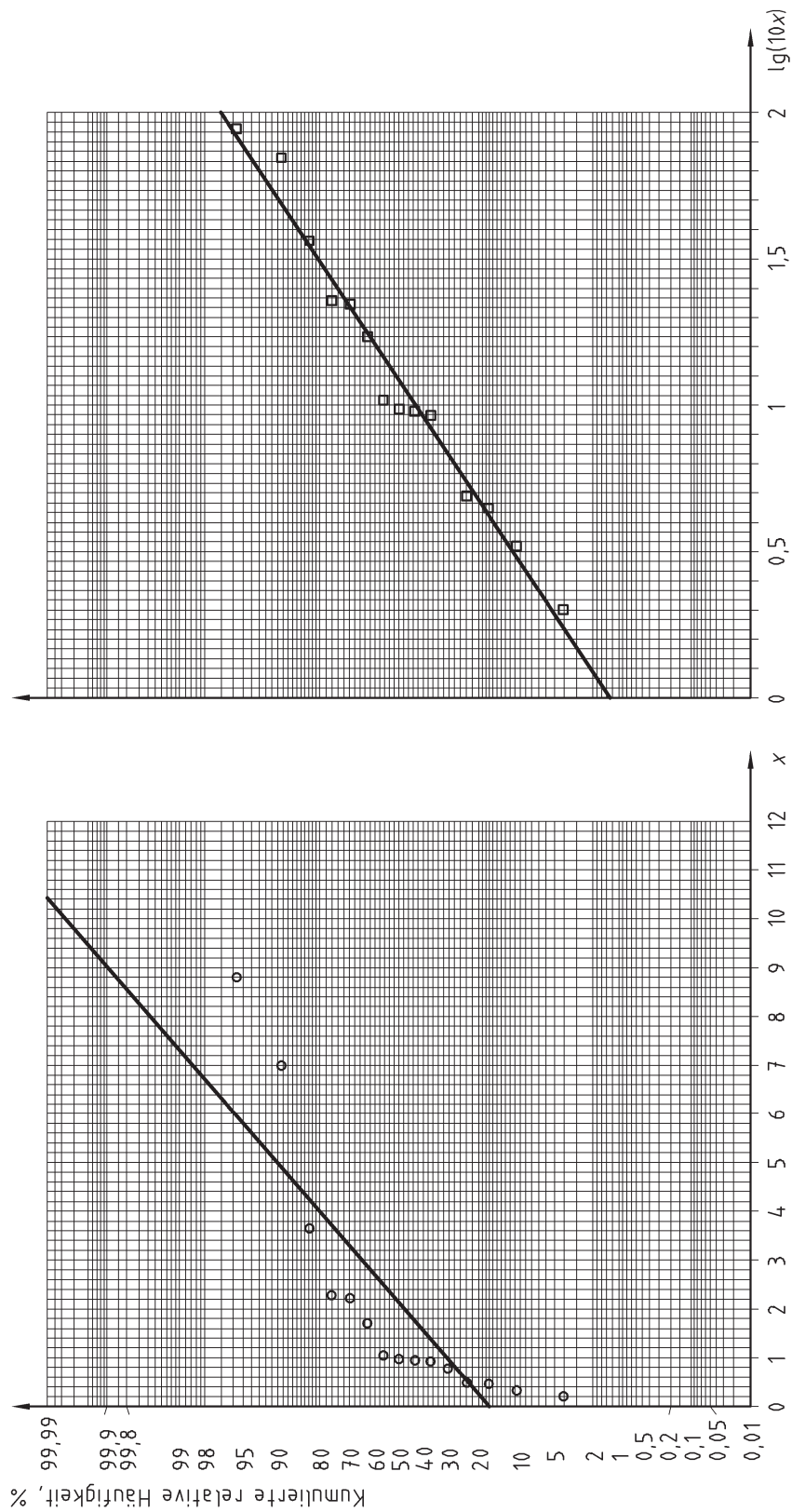


Bild 2 — Grafische Darstellung einer Folge von Messwerten im Wahrscheinlichkeitsnetz

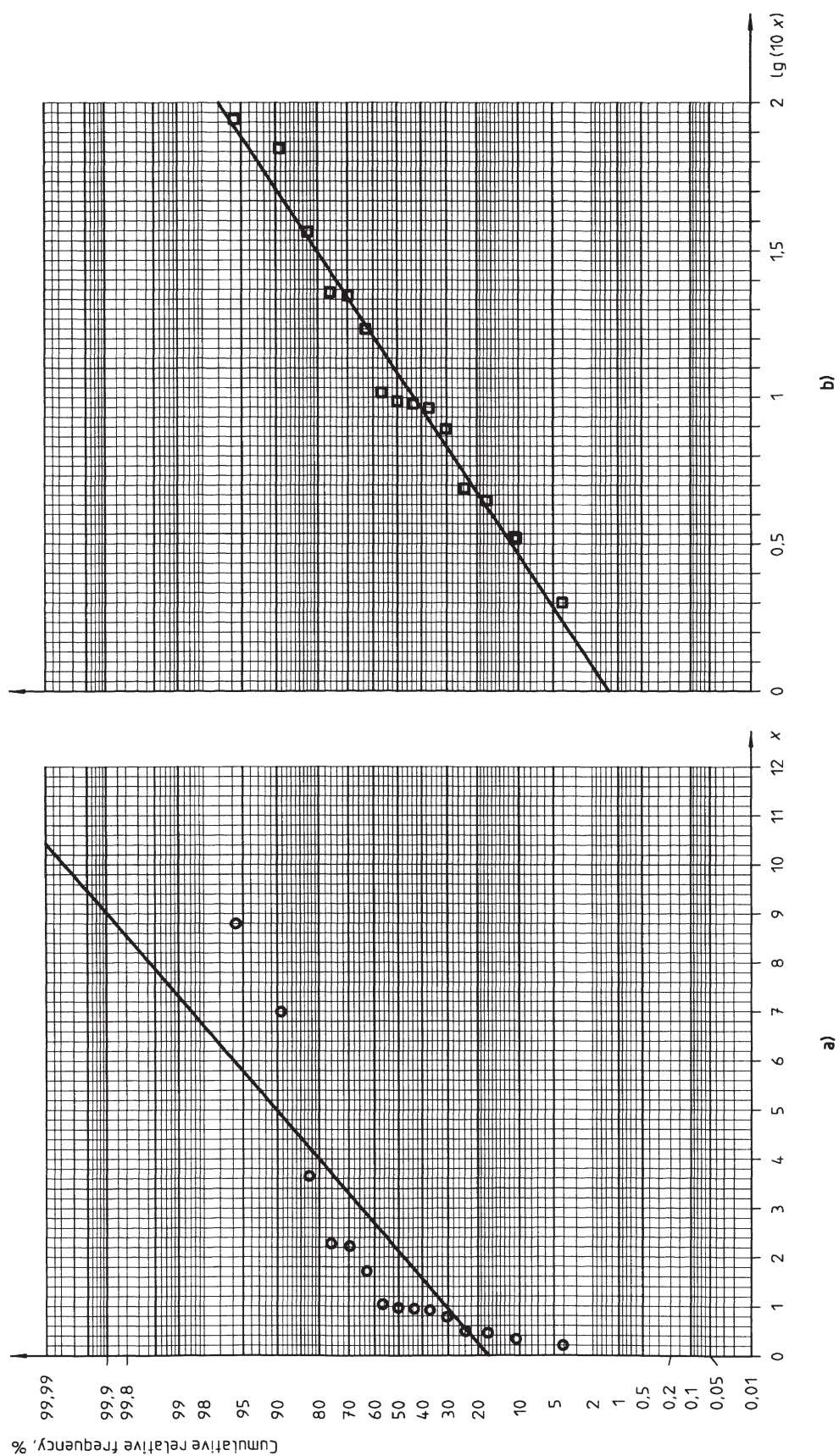


Figure 2 — Graph of a series of observations on normal probability graph paper