

DIN ISO 16269-7

ICS 03.120.30

**Statistische Auswertung von Daten –
Teil 7: Median –
Punktschätzung und Vertrauensbereiche (ISO 16269-7:2001);
Text in Deutsch, Englisch**

Statistical interpretation of data –

Part 7: Median –

Estimation and confidence intervals (ISO 16269-7:2001); Text in German, English

Interprétation statistique des données –

Partie 7: Médiane –

Estimation et intervalles de confiance (ISO 16269-7:2001); Texte en allemand, anglais

Gesamtumfang 27 Seiten

Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ) im DIN

Nationales Vorwort

Dieses Dokument wurde im Technischen Komitee ISO/TC 69, *Anwendung statistischer Verfahren*, vom früheren Unterkomitee 3, *Anwendung statistischer Verfahren in der Normung*, erarbeitet, dessen Sekretariat von AFNOR (Frankreich) gehalten wurde. Das zuständige deutsche Gremium ist der Arbeitsausschuss NA 147-00-02 AA, *Angewandte Statistik*, im Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen.

Anhang A und Anhang B dieses Teils von ISO 16269 dienen nur zur Information. Der NA 147-00-02 AA bevorzugt das im Anhang A dargestellte „klassische Verfahren“, weil dieses im Gegensatz zu der Methode, die im Hauptteil der Norm dargestellt ist, universell für alle Stichprobenumfänge und für alle Vertrauensniveaus anwendbar ist.

Es wird auf die Möglichkeit hingewiesen, dass für Teile dieses Teils von ISO 16269 Patentrechte bestehen können. Das DIN übernimmt keine Verantwortung für die Kennzeichnung solcher Rechte.

Die Normenreihe ISO 16269 umfasst unter dem Haupttitel *Statistische Auswertung von Daten* die folgenden Teile, von denen sich einige noch in Erarbeitung befinden:

- *Teil 1: Leitfaden zur statistischen Auswertung von Daten*
- *Teil 2: Darstellung statistischer Daten*
- *Teil 3: Tests auf Abweichung von der Normalverteilung*
- *Teil 4: Erkennung und Behandlung von Ausreißern*
- *Teil 5: Schätzung und Tests für Mittelwerte und Varianzen für die Normalverteilung, mit Gütefunktionen der Tests*
- *Teil 6: Ermittlung von statistischen Anteilsbereichen*
- *Teil 7: Median — Punktschätzung und Vertrauensbereiche*
- *Teil 8: Ermittlung von Prognosebereichen*

— Leerseite —

Statistische Auswertung von Daten — Teil 7: Median — Punktschätzung und Vertrauensbereiche

Inhalt

	Seite
1 Anwendungsbereich.....	6
2 Normative Verweisungen	6
3 Begriffe und Formelzeichen	6
3.1 Begriffe	6
3.2 Formelzeichen.....	8
4 Anwendbarkeit	8
5 Punktschätzung	8
6 Vertrauensbereich	10
6.1 Allgemeines.....	10
6.2 Klassisches Verfahren	10
6.3 Kleine Stichproben ($5 \leq n \leq 100$).....	10
6.4 Große Stichproben ($n > 100$).....	14
Anhang A (informativ) Klassisches Verfahren zur Ermittlung von Vertrauensgrenzen für den Median	18
Anhang B (informativ) Beispiele	20
 Tabellen	
Tabelle 1 — Genaue Werte von k für Stichprobenumfänge von 5 bis 100: einseitig begrenzter Fall	12
Tabelle 2 — Genaue Werte von k für Stichprobenumfänge von 5 bis 100: zweiseitig begrenzter Fall	13
Tabelle 3 — Werte von u und c für den einseitig begrenzten Fall.....	16
Tabelle 4 — Werte von u und c für den zweiseitig begrenzten Fall.....	16
 Formblätter	
Formblatt A — Berechnung eines Schätzwerts für den Median der Grundgesamtheit	21
Formblatt B — Berechnung eines Vertrauensbereichs für den Median	26

Statistical interpretation of data — Part 7: Median — Estimation and confidence intervals

Contents

	Page
1 Scope	7
2 Normative references	7
3 Terms, definitions and symbols	7
3.1 Terms and definitions	7
3.2 Symbols	9
4 Applicability	9
5 Point estimation	9
6 Confidence interval	11
6.1 General	11
6.2 Classical method	11
6.3 Small samples ($5 \leq n \leq 100$).....	11
6.4 Large samples ($n > 100$)	15
Annex A (informative) Classical method of determining confidence limits for the median	19
Annex B (informative) Examples	21

Tables

Table 1 — Exact values of k for sample sizes varying from 5 to 100: one-sided case.....	12
Table 2 — Exact values of k for sample sizes varying from 5 to 100: two-sided case.....	13
Table 3 — Values of u and c for the two-sided case.....	16
Table 4 — Values of u and c for the two-sided case.....	16

Forms

Form A — Calculation of an estimate of a median	23
Form B — Calculation of a confidence interval for a median	27

1 Anwendungsbereich

Dieser Teil von ISO 16269 legt die Verfahren fest, mit denen ein Punktschätzwert und Vertrauensbereiche für den Median einer beliebigen stetigen Wahrscheinlichkeitsverteilung einer Grundgesamtheit auf der Grundlage einer Zufallsstichprobe aus dieser Grundgesamtheit ermittelt werden. Diese Verfahren sind „verteilungsfrei“, das heißt sie erfordern keine Kenntnis der Familie von Verteilungsfunktionen, zu der die Verteilung der Grundgesamtheit gehört. Ähnliche Verfahren können zur Schätzung von Quartilen und Perzentilen angewendet werden.

ANMERKUNG Der Median ist das zweite Quartil und das fünfzigste Perzentil. Ähnliche Verfahren für andere Quartile oder Perzentile werden in diesem Teil von ISO 16269 nicht beschrieben.

2 Normative Verweisungen

Die folgenden normativen Dokumente enthalten Festlegungen, die durch Verweisung in diesem Text Bestandteil dieses Teils von ISO 16269 sind. Bei datierten Verweisungen, nachfolgenden Ergänzungen oder Revisionen gilt keine dieser neueren Ausgaben. Vertragspartner, deren Vereinbarungen auf diesem Teil von ISO 16269 basieren, werden jedoch gebeten, die Möglichkeit zu prüfen, ob die jeweils neuesten Ausgaben der im Folgenden genannten normativen Dokumente angewendet werden können. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen normativen Dokuments. Die Mitglieder von IEC und ISO führen Verzeichnisse der gegenwärtig gültigen Internationalen Normen.

ISO 2602, *Statistische Auswertung von Ermittlungsergebnissen — Schätzung des Erwartungswertes — Vertrauensbereich*

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

3 Begriffe und Formelzeichen

3.1 Begriffe

Für die Anwendung dieses Teils von ISO 16269 gelten die Begriffe nach ISO 2602 und ISO 3534-1 und die folgenden Begriffe.

3.1.1

***k*-te Ranggröße einer Stichprobe**

Wert des *k*-ten Elements in einer Stichprobe, wenn die Elemente in nicht-absteigender Wertefolge geordnet sind

ANMERKUNG Für eine Stichprobe mit *n* Elementen in nicht-absteigender Wertefolge ist die *k*-te Ranggröße $x_{[k]}$. Dabei ist

$$x_{[1]} \leq x_{[2]} \leq \dots \leq x_{[n]}$$

3.1.2

Median einer stetigen Wahrscheinlichkeitsverteilung

Wert, für den die Anteile der Verteilung zu beiden Seiten des Wertes jeweils genau einer Hälfte entsprechen

ANMERKUNG In diesem Teil von ISO 16269 wird der Median einer stetigen Wahrscheinlichkeitsverteilung Median der Grundgesamtheit genannt und mit dem Symbol *M* bezeichnet.

1 Scope

This part of ISO 16269 specifies the procedures for establishing a point estimate and confidence intervals for the median of any continuous probability distribution of a population, based on a random sample size from the population. These procedures are distribution-free, i.e. they do not require knowledge of the family of distributions to which the population distribution belongs. Similar procedures can be applied to estimate quartiles and percentiles.

NOTE The median is the second quartile and the fiftieth percentile. Similar procedures for other quartiles or percentiles are not described in this part of ISO 16269.

2 Normative references

The following normative documents contain provisions which, through reference in this text, constitute provisions of this part of ISO 16269. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on this part of ISO 16269 are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. Members of ISO and IEC maintain registers of currently valid International Standards.

ISO 2602, *Statistical interpretation of test results — Estimation of the mean — Confidence interval*

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

3 Terms, definitions and symbols

3.1 Terms and definitions

For the purposes of this part of ISO 16269, the terms and definitions given in ISO 2602 and ISO 3534-1 and the following apply.

3.1.1

***k*th order statistic of a sample**

value of the *k*th element in a sample when the elements are arranged in non-decreasing order of their values

NOTE For a sample of *n* elements arranged in non-decreasing order, the *k*th order statistics is $x_{[k]}$ where

$$x_{[1]} \leq x_{[2]} \leq \dots \leq x_{[n]}$$

3.1.2

median of a continuous probability distribution

value such that the proportions of the distribution lying on either side of it are both equal to one half

NOTE In this part of ISO 16269, the median of a continuous probability distribution is called the population median and is denoted by *M*.

3.2 Formelzeichen

a	untere Wertegrenze der Variablen in der Grundgesamtheit
b	obere Wertegrenze der Variablen in der Grundgesamtheit
C	Vertrauensniveau
c	Konstante zur Ermittlung des Wertes von k in Gleichung (1)
k	Rangzahl der für die untere Vertrauensgrenze verwendeten Ranggröße
M	Median der Grundgesamtheit
n	Stichprobenumfang
T_1	aus einer Stichprobe abgeleitete untere Vertrauensgrenze
T_2	aus einer Stichprobe abgeleitete obere Vertrauensgrenze
u	Quantil der standardisierten Normalverteilung
$x_{[i]}$	i -tes kleinstes Element in einer Stichprobe, wenn die Elemente in nicht-absteigender Wertefolge geordnet sind
$\tilde{x}$	Median der Stichprobe
y	mit Hilfe von Gleichung (1) berechneter Zwischenwert zur Ermittlung von k

4 Anwendbarkeit

Das in diesem Teil von ISO 16269 beschriebene Verfahren ist für jede stetige Grundgesamtheit gültig, sofern die Stichprobe zufällig entnommen wird.

ANMERKUNG Kann angenommen werden, dass die Verteilung der Grundgesamtheit näherungsweise eine Normalverteilung ist, so ist der Median der Grundgesamtheit näherungsweise gleich dem Mittelwert der Grundgesamtheit, und die Vertrauensgrenzen sollten in Übereinstimmung mit ISO 2602 berechnet werden.

5 Punktschätzung

Ein Punktschätzwert für den Median der Grundgesamtheit ist durch den Median der Stichprobe, $\tilde{x}$, gegeben. Der Median der Stichprobe ergibt sich durch Nummerierung der Stichprobeneinheiten in nicht-absteigender Wertefolge. Er ist dann

- für ungerade n der Wert mit der Rangzahl $[(n + 1)/2]$ oder
- für gerade n der arithmetische Mittelwert der Werte mit den Rangzahlen $(n/2)$ und $[(n/2) + 1]$.

ANMERKUNG Dieser Schätzwert ist für asymmetrische Verteilungsfunktionen im Allgemeinen nicht erwartungstreu; ein für jegliche Grundgesamtheit erwartungstreuer Schätzwert existiert jedoch nicht.

3.2 Symbols

a	lower bound to the values of the variable in the population
b	upper bound to the values of the variable in the population
C	confidence level
c	constant used for determining the value of k in equation (1)
k	number of the order statistic used for the lower confidence limit
M	population median
n	sample size
T_1	lower confidence limit derived from a sample
T_2	upper confidence limit derived from a sample
u	fractile of the standardized normal distribution
$x_{[i]}$	i th smallest element in a sample when the elements are arranged in a non-decreasing order of their values
$\tilde{x}$	sample median
y	intermediate value calculated to determine k using equation (1)

4 Applicability

The method described in this part of ISO 16269 is valid for any continuous population, provided that the sample is drawn at random.

NOTE If the distribution of the population can be assumed to be approximately normal, the population median is approximately equal to the population mean and the confidence limits should be calculated in accordance with ISO 2602.

5 Point estimation

A point estimate of the population median is given by the sample median, $\tilde{x}$. The sample median is obtained by numbering the sample elements in non-decreasing order of their values and taking the value of

- the $[(n + 1)/2]$ th order statistic, if n is odd, or
- the arithmetic mean of the $(n/2)$ th and $[(n/2) + 1]$ th order statistics, if n is even.

NOTE This estimator is in general biased for asymmetrical distributions, but an estimator that is unbiased for any population does not exist.