

Stochastik

Wahrscheinlichkeitstheorie, Gemeinsame Grundbegriffe
der mathematischen und der beschreibenden Statistik
Begriffe und Zeichen

DIN
13 303
Teil 1

Stochastics; probability theory, common fundamental concepts of mathematical and of descriptive statistics; concepts, signs and symbols

Inhalt

	Seite		Seite
Vorbemerkung	2	2.4 Verteilung spezieller Zufallsvariablen	16
1 Gemeinsame Grundbegriffe der Wahrscheinlichkeitstheorie, der mathematischen und der beschreibenden Statistik	3	3 Parameter	18
1.1 Zufallsvariable	3	3.1 Spezielle Funktionalparameter von Zufallsvariablen und ihren Verteilungen	18
1.2 Funktionen von n -Tupeln von Zufallsvariablen	7	3.2 Arten von Funktionalparametern	20
1.3 Häufigkeiten und Besetzungszahlen	10	3.3 Scharparameter	20
2 Wahrscheinlichkeit	11	4 Stochastische Abhängigkeit und Unabhängigkeit	21
2.1 Grundbegriff Wahrscheinlichkeitsverteilung	11	4.1 Stochastische Unabhängigkeit	21
2.2 Spezielle diskrete Wahrscheinlichkeitsverteilungen	14	4.2 Bedingte Wahrscheinlichkeiten, Dichten und Verteilungsfunktionen	22
2.3 Spezielle Dichten und Verteilungsfunktionen in $\mathbb{R}$	15	4.3 Bedingte Erwartungswerte	24
		Erläuterungen	26
		Stichwortverzeichnis	26

Fortsetzung Seite 2 bis 28

Normenausschuß Einheiten und Formelgrößen (AEF) im DIN Deutsches Institut für Normung e. V.
Ausschuß Qualitätssicherung und angewandte Statistik (AQS) im DIN

Nachdruck, auch auszugsweise, nur mit Genehmigung des DIN Deutsches Institut für Normung e. V., Berlin, gestattet.

Die Normen DIN 13303 Teil 1 und Teil 2 dienen dazu, die Begriffe und Zeichen der Stochastik zu normen, und zwar im vorliegenden Teil 1 die Begriffe und Zeichen der Wahrscheinlichkeitstheorie einschließlich der gemeinsamen Grundbegriffe der mathematischen und der beschreibenden Statistik. Teil 2 über mathematische Statistik liegt z. Z. als Entwurf vor. Ergänzend hierzu behandeln die Normen DIN 55350 Teil 21, Teil 22, Teil 23 (z. Z. Entwurf) und Teil 24 (z. Z. Entwurf) die Begriffe der Statistik aus der Sicht der praktischen Anwendung, wobei dort auf eine strenge mathematische Darstellungsweise im allgemeinen verzichtet wird.

Bezüglich der verwendeten mathematischen Zeichen und Begriffe gelten DIN 1302 und die dort zitierten weiteren Normen mathematischen Inhalts, insbesondere DIN 5473.

Als Zeichen für die Teilmengenrelation (siehe DIN 5473, Ausgabe Juni 1976, Nr. 1.6) wird in der Stochastik $\subset$ benutzt, das Komplement von A (siehe DIN 5473, Ausgabe Juni 1976, Nr. 1.11) wird mit $\bar{A}$ oder A^c bezeichnet. Ferner wird beim Mengenbildungoperator (siehe DIN 5473, Ausgabe Juni 1976, Nr. 1.4) statt $\{x|\varphi\}$ in dieser Norm $\{x:\varphi\}$ geschrieben, um Verwechslungen mit der Bezeichnung für bedingte Wahrscheinlichkeiten zu vermeiden. Es ist in der Stochastik üblich, diese Schreibweise noch weiter zu kürzen:

Statt $\{\omega:X(\omega)\leq x\}$ wird $\{X(\omega)\leq x\}$ oder $\{X\leq x\}$,

statt $\{\omega:X(\omega)\in A\}$ wird kurz $\{X\in A\}$,

statt $\{\omega:X(\omega)=x\}$ und $\{Y(\omega)=y\}$ wird kurz $\{X=x, Y=y\}$ geschrieben. Ferner wird in der Stochastik oft eine disjunkte Vereinigung gebraucht, die mit dem Summenzeichen geschrieben wird. Wenn $A_i \cap A_j = \emptyset$ für $i \neq j$, so

schreibt man $\sum_{i=1}^{\infty} A_i$ für $\bigcup_{i=1}^{\infty} A_i$, ebenso $A+B$ für $A \cup B$,

wenn $A \cap B = \emptyset$; entsprechend werden $\sum_{i=1}^n A_i$ und $\sum_{i \in I} A_i$ definiert.

Es wurde versucht, eine gemeinsame Sprache für alle Statistiker, und zwar für Mathematiker und für Anwender, zu finden.

Die Anmerkung 1 zu 1.1.13 empfiehlt, den Begriff der Stichprobe nur für die reale Entnahme von Stichproben zu verwenden. Ebenso sollten die Begriffe Gesamtheit und (Auswahl-)Einheit nur im Zusammenhang mit der realen Stichprobentnahme verwendet werden. Zum Begriff der Einheit im statistischen Sinne, der Gesamtheit und der Stichprobe siehe auch DIN 55350 Teil 14 (z. Z. Entwurf).

Inbesondere wird zwischen den Parametern einer theoretischen Wahrscheinlichkeitsverteilung und den Kennwerten der (bei der Stichprobentnahme) ermittelten Meßreihen unterschieden (siehe Anmerkungen 1 und 3 zu 1.2 sowie zu 3.1). Zum Begriff des Kennwertes siehe auch DIN 55302 Teil 1.

Für die bei der Ermittlung von Daten in realen Experimenten zu verwendenden Begriffe Messen, Zählen, Prüfen, Meßgröße, Meßwert, Meßreihe, Meßergebnis (siehe Anmerkung zu 1.1) siehe DIN 1319 Teil 1 und DIN 1319 Teil 3. Für die Begriffe „qualitatives Merkmal“ und „Merkmalswert“ siehe DIN 55350 Teil 12.

Für den in der Definition der Zufallsgröße (siehe Nr. 1.1.6 und Anmerkung zu 1.1.6) verwendeten Begriff „(physikalische) Größe“ siehe DIN 1313.

Die internationale Terminologie wurde berücksichtigt, insbesondere die von der International Organization for Standardization (ISO) herausgegebene Internationale Norm ISO 3534 „Statistics – Vocabulary and Symbols“. Abweichungen von den Zeichen dieser Norm werden in den Anmerkungen erläutert.

Die Abschnitte dieser Norm sind in der Regel so in Tabellenform aufgebaut, daß die mathematisch weniger anspruchsvollen Begriffe vor den nur dem Spezialisten verständlichen Begriffen erscheinen, soweit dies die Systematik der Abschnitte zuläßt. In diesem Sinne erscheint in Abschnitt 1.1 eine eigene Spalte „Definition für endliche oder abzählbar unendliche Ergebnismengen“, um die Definition nicht mit maßtheoretischen Details zu belasten. Die Namen der Begriffe erscheinen in den meisten Abschnitten vor den Zeichen für diese, da die Normung der Namen für die Verständigung im Bereich der Stochastik, insbesondere Statistik, grundlegend ist. In den Abschnitten 1.2, 3.1, 4.2 und 4.3 erscheinen die Zeichen vor den Namen der Begriffe, da dort die Normung der Zeichen größere Bedeutung hat. Bei den Zeichen wird deutlich gesagt, wann diese durch die vorliegende Norm empfohlen werden, z. B. im Kopf von Spalte 2 in Abschnitt 3.1 und im Vortext von Abschnitt 4.2. Wenn die verwendeten Zeichen nicht ausdrücklich empfohlen werden, können diese durch andere Zeichen ersetzt werden. Darauf wird durch Fußnoten gelegentlich hingewiesen, z. B. in den Abschnitten 1.1 und 2.1.

1 Gemeinsame Grundbegriffe der Wahrscheinlichkeitstheorie, der mathematischen und der beschreibenden Statistik

1.1 Zufallsvariable

Siehe Anmerkung zu 1.1 für die Anwendung der in Spalte 2 festgelegten Begriffe in der beschreibenden Statistik (im Sinne der Beschreibung von in realen Experimenten erhobenen Daten). Zur Bedeutung der Spalte 5 siehe Anmerkung zu Spalte 5 von 1.1.

1	2	3	4	5	6
Nr	Name	Zeichen ¹⁾	Definition		Bemerkung
			für endliche oder abzählbar unendliche Ergebnismengen	Zusatzvoraussetzungen für überabzählbare Ergebnismengen	
1.1.1	Ergebnis	ω	Grundbegriff (mit der Bedeutung: mögliches Ergebnis bei dem durchzuführenden, zu beschreibenden Experiment)		Zur Wahl des Wortes „Ergebnis“ siehe Anmerkung zu 1.1.1.
1.1.2	Ergebnismenge	Ω	Menge aller möglichen Ergebnisse		„Stichprobenraum“ wird in der Literatur oft anstelle der hier empfohlenen Benennung „Ergebnismenge“ verwendet.
1.1.3	Meßraum	$(\Omega, \mathcal{A})$	ein Paar aus einer nicht leeren Menge Ω und einer Menge $\mathcal{A}$ von Teilmengen von Ω , wobei $\mathcal{A}$ die Menge aller Teilmengen von Ω ist oder mindestens eine σ -Algebra	$\mathcal{A}$ ist eine σ -Algebra in Ω .	Zum Begriff des Maßraumes siehe Nr 2.1.2, Spalte 6. Eine σ -Algebra $\mathcal{A}$ in Ω ist eine nicht leere Menge von Teilmengen von Ω mit: $\emptyset \in \mathcal{A}$; $\Omega \in \mathcal{A}$; wenn $A \in \mathcal{A}$, so $\bar{A} \in \mathcal{A}$; wenn $A_i \in \mathcal{A}$, so $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.
1.1.4	Ereignis	A auch $B, C, \dots$	Teilmenge von Ω aus $\mathcal{A}$		
1.1.5	Zufallsvariable, $\mathcal{X}$ -wertige Zufallsvariable	X auch $Y, Z, \dots$	eine Funktion mit dem Definitionsbereich Ω , deren Werte einer Menge $\mathcal{X}$ angehören	eine Funktion $\langle \omega \mapsto X(\omega) \rangle$ mit $X: (\Omega, \mathcal{A}) \rightarrow (\mathcal{X}, \mathcal{B})$, d. h. für alle $B \in \mathcal{B}$ gilt: $\{\omega : X(\omega) \in B\} \in \mathcal{A}$, wobei $(\Omega, \mathcal{A})$, $(\mathcal{X}, \mathcal{B})$ Meßräume sind	Zum Gebrauch des Wortes „Zufallsvariable“ im Vergleich zu „Zufallsgröße“ usw. siehe Anmerkung 1 zu 1.1.5 und Anmerkung zu 1.1.6. Zur Bezeichnung des „Bild“-Meßraumes $(\mathcal{X}, \mathcal{B})$ siehe Anmerkung 2 zu 1.1.5.
1.1.5.1	diskrete Zufallsvariable		jede Zufallsvariable	Es gibt eine endliche oder abzählbar unendliche Menge $T' \in \mathcal{B}$ mit $P\{X \in T'\} = 1$.	Für die Definition des Wahrscheinlichkeitsmaßes P siehe Nr 2.1.1.

¹⁾ Die in dieser Spalte verwendeten Zeichen (mit Ausnahme des Zeichens I für Indikatoren) können durch andere Zeichen ersetzt werden, insbesondere wenn diese informativer sind, z. B. ω durch $(x_1, \dots, x_n)$, wenn das Ergebnis aus den Meßwerten $x_1, \dots, x_n$ besteht.

1	2	3	4	5	6
Nr	Name	Zeichen ¹⁾	Definition		Bemerkung
			für endliche oder abzählbar unendliche Ergebnismengen	Zusatzvoraussetzungen für überabzählbare Ergebnismengen	
1.1.5.2	ganzzahlige Zufallsvariable		Zufallsvariable, deren Werte ganze Zahlen sind	Für alle $k \in \mathbb{Z}$ gilt $\{\omega : X(\omega) = k\} \in \mathcal{A}$.	
1.1.5.3	reellwertige Zufallsvariable		Zufallsvariable, deren Werte reelle Zahlen sind	Für alle $x \in \mathbb{R}$ gilt $\{\omega : X(\omega) \leq x\} \in \mathcal{A}$.	
1.1.6	Zufallsgröße		Zufallsvariable, deren Werte Werte einer Größe sind		Bezüglich des Begriffs einer (physikalischen) Größe siehe DIN 1313. Siehe Anmerkung zu 1.1.6.
1.1.7	Zufallsvektor	$(X_1, \dots, X_n)$ auch $\mathbf{X}, \mathbf{X}$	n -Tupel von Zufallsgrößen $X_1, \dots, X_n$		
1.1.8	Zufallsfolge	$(X_n)_{n \in \mathbb{N}}$ auch $X_{\mathbb{N}}$	Folge von Zufallsvariablen $X_0, X_1, \dots$		Anstelle von $\mathbb{N}$ sind auch andere endliche oder abzählbar unendliche Indexmengen zulässig, z.B. $\mathbb{Z}$ für die Folge $\dots, X_{-1}, X_0, X_1, \dots$
1.1.9	Zufallsfunktion		Zufallsvariable, deren Werte Funktionen $\langle t \mapsto X_t \rangle$ sind		Zum Begriff des stochastischen Prozesses siehe Anmerkung zu 1.1.9.
1.1.10	Indikator des Ereignisses A	I_A	diejenige Zufallsvariable, die für $\omega \in A$ den Wert 1 und sonst den Wert 0 annimmt		Die außerhalb der Stochastik übliche Bezeichnung „charakteristische Funktion der Menge A “ ist wegen Nr 3.1.17 nicht zulässig. Siehe auch Anmerkung zu 1.1.10.
1.1.11	Zufallsziffern		ermittelter Wert einer Folge von über den Ziffern 0, 1, ..., 9 gleichverteilten stochastisch unabhängigen Zufallsvariablen		Zum Gebrauch der Worte „ermittelter Wert“ siehe Beschreibung B in der Anmerkung zu 1.1. Für die Definition der Gleichverteilung und der stochastischen Unabhängigkeit siehe Nr 2.2.1 und Nr 4.1.2.2. Bei anderen Ziffernmengen ist diese anzugeben. Zu Pseudozufallsziffern und -zahlen siehe Anmerkungen zu 1.1.11 und 1.1.12.
1) Siehe Seite 3					

Fortsetzung des Abschnitts 1.1

1	2	3	4	5	6
Nr	Name	Zeichen ¹⁾	Definition		Bemerkung
			für endliche oder abzählbar unendliche Ergebnismengen	Zusatzvoraussetzungen für überabzählbare Ergebnismengen	
1.1.12	(rechteckverteilte) Zufallszahl			ermittelter Wert einer Folge von über dem Intervall (0, 1) rechteckverteilten stochastisch unabhängigen Zufallsvariablen	Für die Definition der Rechteckverteilung siehe Nr 2.3.1. Bei anders verteilten Zufallszahlen ist die Verteilung anzugeben.
1.1.13	Zufallsstichprobe		ermittelter Wert (Realisation) einer über einer endlichen Menge von möglichen Stichproben gleichverteilten Zufallsvariable		Zufallsstichprobe soll nicht im Sinne eines n -Tupels von Zufallsvariablen benutzt werden. Siehe Anmerkung 2 zu 1.1.13.
1.1.14	zentrierte Zufallsvariable		reellwertige Zufallsvariable X mit $E X = 0$, falls $E X$ existiert		Für die Definition des Erwartungswertes $E X$, des Medians von X und der Varianz $Var X$ siehe Nr 3.1.3, Nr 3.1.1.1 und Nr 3.1.4.
1.1.14.1	an einem Median zentrierte Zufallsvariable		reellwertige Zufallsvariable mit 0 als Median		
1.1.15	standardisierte Zufallsvariable		reellwertige Zufallsvariable X mit $E X = 0$ und $Var X = 1$		
1) Siehe Seite 3					

Anmerkungen**Zu 1.1**

Mit den in Abschnitt 1.1 genannten Begriffen der Stochastik können auf zwei Weisen die Daten beschrieben werden, die in einem realen Experiment ermittelt werden sollen. Reale Experimente sind z. B.:

- Messen,
 - Zählen,
 - Prüfen,
 - Bestimmung eines qualitativen Merkmals,
 - Bestimmung des Ablaufes eines Prozesses.
- Zur Definition der Begriffe a) bis c) siehe DIN 1319 Teil 1, zu d) siehe DIN 55 350 Teil 12.

Beschreibung A

Bei der einmaligen Durchführung eines realen Experimentes wird in den genannten Fällen a) bis e) als Ergebnis ermittelt:

- ein Meßwert,
- eine Zahl,
- eine Feststellung,
- ein Merkmalswert (definiert als Ausprägung des Merkmals),
- ein Pfad durch die Menge aller möglichen Zustände.

Bei der n -maligen Durchführung des gleichen realen Experimentes ist dann das Ergebnis ein n -Tupel $\omega = (x_1, \dots, x_n)$ a) von Meßwerten, b) von Zahlen usw.

Braucht man wie in Abschnitt 1.3 einen Oberbegriff für die soeben Meßwerte, Zahlen, Merkmalswerte usw. genannten Komponenten $x_1, \dots, x_n$ des Ergebnisses ω , so kann man von den Einzelergebnissen $x_1, \dots, x_n$ sprechen und von der Menge $\mathcal{X}$ aller möglichen Einzelergebnisse. Beachte $\Omega = \mathcal{X}^n$.

Bei mehrstufigen Experimenten ist das Ergebnis eine endliche oder unendliche Folge $\omega = (x_1, x_2, \dots)$ a) von Meßwerten, b) von Zahlen usw., wobei die Meßwerte $x_1, x_2, \dots$ in der 1., 2., ... Stufe des Experimentes gemessen werden.

Dabei können die einzelnen Stufen verschiedenartige reale Experimente sein.

Beschreibung B

Bei der einmaligen Durchführung eines realen Experimentes wird ein Wert einer Zufallsvariablen ermittelt, und zwar genauer in den oben im ersten Absatz dieser Anmerkung genannten Fällen a) bis e):

- ein Wert einer Zufallsgröße, nämlich der Meßwert einer Meßgröße,
- ein Wert einer Zufallsvariablen, nämlich eine Zahl,
- ein Wert einer Zufallsvariablen, deren Werte Feststellungen sind,
- ein Wert einer Zufallsvariablen, deren Werte Ausprägungen eines Merkmals sind,
- ein Wert einer Zufallsfunktion, nämlich eine Funktion $(t \rightarrow x_t)$.

Dabei werden die möglichen Ergebnisse und die Ergebnismenge Q in der Regel nicht angegeben.

Bei der n -maligen Durchführung des gleichen Experimentes wird der Wert eines n -Tupels von Zufallsvariablen ermittelt, deren Werte in den n einzelnen Experimenten ermittelt werden.

Bei mehrstufigen Experimenten wird der Wert einer Zufallsfolge (X_i) ermittelt. Dabei wird der Wert der ersten Zufallsvariablen X_1 in der 1. Stufe, der Wert von X_2 in der 2. Stufe usw. ermittelt.

Die Beschreibung A ist begrifflich einfacher als die Beschreibung B und wird daher u. a. im Schulunterricht bei der Einführung in die Wahrscheinlichkeitsrechnung unter Verwendung der Beispiele a), b) und d) benutzt. Erst bei der Berechnung des Mittelwertes, der empirischen Standardabweichung usw. treten Funktionen des Ergebnisses auf, für die die unten in Spalte 2 des Abschnitts 1.2 empfohlenen Zeichen auch im Falle der Beschreibung A zu verwenden sind.

Die Beschreibung B hat den Vorteil, bei allen Größen zwischen der Zufallsgröße und ihrem ermittelten Wert zu unterscheiden, und zwar unabhängig davon, ob es sich um die unmittelbar zu messenden Werte von Größen handelt oder um Werte, die erst aus den ermittelten Werten unmittelbar messbarer Größen zu berechnen sind.

Zu Spalte 5 von 1.1

Zusatzvoraussetzungen für überabzählbare Ergebnismengen sind aus einem innermathematischen Grund notwendig: Es gibt keine nicht trivialen Wahrscheinlichkeitsmaße $A \mapsto P(A)$, so daß $P(A)$ für alle $A \subset \mathbb{R}$ so definiert werden kann, daß die in Nr 2.1.1 genannten Eigenschaften a) bis c) erfüllt sind.

Zu 1.1.1

Das Wort „Ergebnis“ wurde trotz der Verwechslungsgefahr mit anderen Bedeutungen dieses Wortes gewählt: a) Ergebnis im Sinne von Ergebnis einer Rechnung, b) das aus den statistischen Daten zu berechnende Ergebnis (siehe z. B. DIN 55303 Teil 2, Entwurf Mai 1978, Formblatt A, A', ..., c) Ergebnis als Kurzform für Meßergebnis im Sinne von DIN 1319 Teil 1.

Alle anderen in der Wahrscheinlichkeitstheorie und deren Anwendungen benutzten Wörter sind bereits anders definiert (Stichprobe im Sinne der Entnahme einer realen Probe, Ausfall im Sinne des Ausfallens einer Maschine, Ausgang im Sinne der Begriffspaare Eingangs- und Ausgangsgrößen (siehe z. B. DIN 40148 Teil 1 und DIN 40146

Teil 1)). Das Wort „Resultat“ ist sprachlich ein Synonym für Ergebnis.

Zu 1.1.5

1 Das Wort „Variable“ in „Zufallsvariable“ ist als Leerstelle für genauer spezifizierte Begriffe anzusehen. Dementsprechend wird empfohlen, das Wort Variable in Zufallsvariable durch die Namen dieser Begriffe wie Größe, Vektor, Funktion, Folge, Matrix usw. zu ersetzen, siehe z. B. Nr 1.1.6 bis Nr 1.1.9. Ausnahmen von der Regel sind die in Nr 1.1.11 bis Nr 1.1.13 definierten Begriffe der Zufallsgrößen, Zufallszahl und Zufallsstichprobe, die sich auf einen Wert einer entsprechenden Zufallsvariablen beziehen.

2 In Abhängigkeit von der Struktur des „Bild“-Meßraumes werden in der Literatur verschiedene Bezeichnungen verwendet, z. B. $(Q, \mathcal{Q}')$, wenn keine Struktur vorausgesetzt wird $(\mathcal{X}, \mathcal{B})$, wenn $\mathcal{X}$ ein topologischer Raum (siehe DIN 13302, Ausgabe Juni 1978, Nr 9.1) ist und $\mathcal{B}$ die kleinste σ -Algebra ist, die alle offenen Mengen enthält. Da der letztgenannte Fall sehr häufig ist und $(\mathcal{X}, \mathcal{B})$ sich deutlich von $(Q, \mathcal{Q}')$ typographisch unterscheidet, wird in dieser Norm die Bezeichnung $(\mathcal{X}, \mathcal{B})$ verwendet.

Zu 1.1.6

In der Wahrscheinlichkeitstheorie und mathematischen Statistik werden zur Zeit die Doppelwörter „Zufallsvariable“ und „Zufallsgröße“ oft als Synonyme behandelt. Es wird jedoch vornehmlich empfohlen, das Wort Zufallsgröße nur dann zu verwenden, wenn die Werte der Zufallsgröße Werte einer (physikalischen) Größe sind.

Zu 1.1.9

Der Name „Zufallsfunktion“ für die Abbildungen $\langle \omega \mapsto \langle t \mapsto X_t(\omega) \rangle | T \rangle$ wurde aus den aus der Anmerkung 1 zu 1.1.5 ersichtlichen systematischen Gründen hier angegeben. Weit verbreitet ist die Benennung „stochastischer Prozeß“, wobei die angegebene Abbildung meist in der Gestalt einer Familie von Zufallsvariablen $(X_t)_{t \in T}$ angegeben wird, kurz X_T , ganz ausführlich $\langle \omega \mapsto (X_t(\omega))_{t \in T} \rangle$. Die mit dem Gebiet der stochastischen Prozesse mit nicht abzählbarem Definitionsbereich T zusammenhängenden Begriffe, Benennungen und Bezeichnungsweisen werden in dieser Norm nicht behandelt.

Zu 1.1.10

Die Funktion $\langle \omega \mapsto I_A(\omega) \rangle$ wird in der Literatur auch Indikatorfunktion genannt.

Zu 1.1.11 und 1.1.12

Pseudozufallsziffern oder -zahlen werden mit einem algebraischen Algorithmus erzeugt und wie Zufallsziffern oder -zahlen verwendet.

Zu 1.1.13

1 Die Bezeichnung „Stichprobe“ sollte nur für die Entnahme realer Stichproben, etwa aus einer Menge Schrauben, benutzt werden. Daran anschließend wird in der Literatur die Meßreihe der Meßwerte der Länge der Schrauben aus der realen Stichprobe auch als Stichprobe bezeichnet.

2 Ein n -Tupel von Zufallsvariablen ist u. a. als Modell für die folgenden zwei begrifflich zu unterscheidenden Situationen der Stichprobenentnahme aus einer Gesamtheit von (Auswahl-)Einheiten geeignet.

- An n systematisch oder
- an n zufällig ausgewählten Einheiten wird dieselbe Größe gemessen (z. B. eine Reaktionszeit), deren n Meßwerte die ermittelten Werte von n Zufallsgrößen sind.

In beiden Fällen ist der ermittelte Wert des n -Tupels von Zufallsvariablen in der Regel keine Zufallsstichprobe im Sinne von Nr 1.1.13. Im Fall b) wird in der Literatur jedoch auch von Zufallsstichproben gesprochen.

1.2 Funktionen von n -Tupeln von Zufallsvariablen

Es liege die Realisation $(x_1, \dots, x_n)$ eines n -Tupels $(X_1, \dots, X_n)$ von Zufallsvariablen vor. Anstelle von Realisation sind auch andere Benennungen für $(x_1, \dots, x_n)$ üblich: Meßreihe (siehe DIN 1319 Teil 3), die aus den ermittelten Werten $x_1, \dots, x_n$ besteht; Stichprobe (siehe Anmerkung 1 zu 1.1.13), die aus den einzelnen Stichprobenwerten $x_1, \dots, x_n$ besteht. Zur Wahl der kleinen und großen lateinischen Buchstaben siehe Anmerkung 2 zu 1.2. Zum Begriff des Kennwertes siehe Anmerkung 1 zu 1.2.

Voraussetzung für die Anwendung der Nr 1.2.1 bis Nr 1.2.7 ist, daß die Werte der Zufallsvariablen (linear) geordnet werden können. Voraussetzung für die Anwendung der Nr 1.2.4 (soweit dort eine Summe von x -Werten auftritt), der Nr 1.2.5 und ab Nr 1.2.8 ist, daß Zufallsvariablen vorliegen, bei denen insbesondere die vorkommenden Additionen und Subtraktionen sinnvoll sind.

1	2	3	4	5	6	7
Nr	empfohlenes Zeichen	Verwendung als Zeichen für		Name	Definition	Bemerkung
		die ermittelten Werte	entsprechende Zufallsvariablen			
1.2.1	()	$x_{(1)}, \dots, x_{(n)}$ auch: $x_{(1)n}, \dots, x_{(n)n}$	$X_{(1)}, \dots, X_{(n)}$ auch: $X_{(1)n}, \dots, X_{(n)n}$	Ordnungsstatistik	das geordnete n -Tupel der ermittelten Werte $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$	
1.2.2	r oder s bzw. R oder S	$r_1, \dots, r_n$ auch: $r_1, n, \dots, r_n, n$	$R_1, \dots, R_n$ auch: $R_1, n, \dots, R_n, n$	Rangzahlen	Sind alle Werte der Zufallsvariablen verschieden, so gilt: $r_i = k$, wenn $x_i = x_{(k)}$	Sind nicht alle Werte der Zufallsvariablen verschieden, so gibt es verschiedene Definitionen.
1.2.3	$F, G, H, \dots$	$\langle x \rightarrow F_n(x) \rangle$ Statt F_n auch $F_{n,x}$	$\langle x \rightarrow F_n(x) \rangle$ Statt F_n auch $F_{n,x}$	(empirische) Verteilungsfunktion	$F_n(x) = \frac{1}{n} \text{card} \{j : x_j \leq x\}$	Zum Gebrauch des Wortes „empirisch“ siehe Anmerkung 3 zu 1.2
1.2.4	$\sim$	$\tilde{x}$	$\tilde{X}$	(empirischer) Median	für n ungerade: $\tilde{x} = x_{((n+1)/2)}$; für n gerade: $\tilde{x} = \frac{1}{2} (x_{(n/2)} + x_{(n/2+1)})$, in diesem Fall sind auch andere Definitionen zulässig (und bei Ordinalskalenwerten ²⁾ sogar notwendig), solange gilt: $x_{(n/2)} \leq \tilde{x} \leq x_{(n/2+1)}$	Empirische p -Quantile und deren Funktionen werden entsprechend den in Nr 3.1.1 definierten theoretischen p -Quantilen definiert, indem man anstelle der Wahrscheinlichkeitsverteilung P die Häufigkeitsverteilung h_n benutzt.
1.2.5.1 1.2.5.2				Spannweite Quasispannweite	$x_{(n)} - x_{(1)}$ $x_{(n-i)} - x_{(i+1)}$ mit $i = 1, 2, \dots$	
1.2.6	r_s	$r_{s,xy}$	$r_{s,XY}$ oder R	Spearman'scher Rangkorrelationskoeffizient	$r_{s,xy} = 1 - \frac{6}{n^3 - n} \sum_{i=1}^n (r_i - s_i)^2$	Gewöhnlicher Korrelationskoeffizient (siehe Nr 1.2.13) r_{rs} für die Rangzahlen $r_1, \dots, r_n$ und $s_1, \dots, s_n$ der Werte $(x_1, \dots, x_n)$ bzw. $(y_1, \dots, y_n)$ der n -Tupel von Zufallsvariablen.
²⁾ Siehe z. B. DIN 53804 Teil 3.						

1	2	3	4	5	6	7
Nr	empfohlenes Zeichen	Verwendung als Zeichen für		Name	Definition	Bemerkung
		die ermittelten Werte	entsprechende Zufallsvariablen			
1.2.7	r_K	$r_{K,xy}$	$r_{K,xy}$ oder T	Kendallscher Rangkorrelationskoeffizient	$r_{K,xy} = \frac{1}{\binom{n}{2}} \sum_{i < j} \text{sgn}(x_i - x_j) \cdot \text{sgn}(y_i - y_j)$	$\text{sgn}(x_i - x_j) = \text{def} \begin{cases} 1, & \text{wenn } x_i > x_j \\ 0, & \text{wenn } x_i = x_j \\ -1, & \text{wenn } x_i < x_j \end{cases}$
1.2.8	—	$\bar{x}$	$\bar{X}$	Mittelwert, arithmetisches Mittel	$\bar{x} = \frac{1}{n} (x_1 + \dots + x_n)$	Siehe auch Nr 1.2.17.
1.2.9	$m'_r(c)$ m'_r für $m'_r(0)$ m_r für $m'_r(\bar{x})$			(empirisches) Moment der Ordnung r bzgl. c	$m'_r(c) = \frac{1}{n} \sum_{i=1}^n (x_i - c)^r$	Im Fall $c = 0$ läßt man den Zusatz „bzgl. 0“ weg. Im Fall $c = \bar{x}$ spricht man vom zentralen Moment.
1.2.9.1	m_2			(empirisches) zweites zentrales Moment	$m_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$	
1.2.10	s^2	s^2 oder s_x^2	s^2 oder s_X^2	(empirische) Varianz	$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$	Siehe Anmerkung zu 1.2.10 und 1.2.12. Es sind auch andere Indizierungen von s^2 zulässig.
1.2.10.1	s	s oder s_x	s oder s_X	(empirische) Standardabweichung	$s_x = \sqrt{s_x^2}$	Für eine Verwendung des Zeichens S siehe Nr 2.4.10.
1.2.11	v	v oder v_x	v oder v_X	(empirischer) Variationskoeffizient	$v_x = \frac{s_x}{\bar{x}}$ mit $x_i \geq 0$	
1.2.12	s	s_{xy}	s_{XY}	(empirische) Kovarianz	$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$	Siehe Anmerkung zu 1.2.10 und 1.2.12.
1.2.12.1				(empirische) Kovarianzmatrix	(s_{ij}) mit $s_{ii} = s_{x_i}^2$ und $s_{ij} = s_{x_i x_j}$	Es liegen die Realisationen $x_1, \dots, x_n$ von n Zufallsvektoren $X_1, X_2, \dots, X_n$ mit $x_i = (x_{i1}, \dots, x_{im})'$ vor.
1.2.13	r	r_{xy}	r_{XY}	(empirischer) Korrelationskoeffizient	$r_{xy} = \frac{s_{xy}}{s_x s_y}$	
1.2.14	b	b_{yx}	b_{YX}	(empirischer) Regressionskoeffizient von y auf x	$b_{yx} = \frac{s_{xy}}{s_x^2} = \frac{r_{xy} s_y}{s_x}$	$y = a_{yx} + b_{yx}x$ ist die Gleichung der empirischen Regressionsgeraden von y auf x . Vertauschung von x mit y liefert die Regressionsgerade von x auf y .
1.2.15	a	a_{yx}	a_{YX}	(empirische) Regressionskonstante von y auf x	$a_{yx} = \bar{y} - b_{yx} \bar{x}$	

Fortsetzung des Abschnitts 1.2

1	2	3	4	5	6	7
Nr	empfohlenes Zeichen	Verwendung als Zeichen für		Name	Definition	Bemerkung
		die ermittelten Werte	entsprechende Zufallsvariablen			
1.2.16	.	$x_{i,k}$	$X_{i,k}$		$x_{i,k} = \sum_{j=1}^I \sum_{l=1}^L x_{ijkl}$ $\bar{x}_{ij,l} = \frac{1}{K} x_{ij,l} = \frac{1}{K} \sum_{k=1}^K x_{ijkl}$ – bedeutet Mittelung über alle punktierten Indizes.	Die in Nr 1.2.16 empfohlene Punktbezeichnung für Summen und die in Nr 1.2.17 empfohlene Punktbezeichnung mit Kopfstich für arithmetische Mittel dient der Vereinheitlichung derselben in den Bereichen der Varianzanalyse und Kontingenztafeln, siehe Beispiel zu Abschnitt 1.3.
1.2.17	-	$\bar{x}_{ij,l}$ (nicht: $x_{ij,l}$)	$\bar{X}_{ij,l}$			
1.2.18	SQ			Quadratsummen (z.B. der Varianzanalyse)	z.B.: $SQ_b = I \sum_{i=1}^I (\bar{x}_{i.} - \bar{x}_{..})^2$	(x_{ij}) ist eine (I, J) -Matrix.

Anmerkungen**Zu 1.2**

1 Die in einem realen Experiment ermittelten Werte der in Nr 1.2.4 bis Nr 1.2.15 aufgeführten Funktionen werden auch Kennwerte der Stichprobe (siehe z.B. DIN 55302 Teil 1) genannt.

Wird die Stichprobe aus einer realen Gesamtheit von möglichen Auswahleinheiten gezogen, so sind die Kennwerte der Stichprobe von den entsprechenden Kennwerten der Gesamtheit zu unterscheiden.

Der Begriff des Kennwertes einer Gesamtheit fällt unter den Oberbegriff des Parameters einer Verteilung, siehe dazu Abschnitt 3.

2 **Zur Wahl der Zeichen in Spalte 3 und 4:** Als Zeichen für Zufallsvariablen werden in der Regel große lateinische Buchstaben verwendet. Der entsprechende kleine lateinische Buchstabe bezeichnet dann den ermit-

telten Wert (die Realisation) dieser Zufallsvariablen. Werden Zufallsvariablen wie in Nr 1.2.6 und Nr 1.2.7 sowie Nr 1.2.9 bis Nr 1.2.15 abweichend von dieser Regel mit kleinen lateinischen Buchstaben bezeichnet, so wird empfohlen, diese Regel auf die Indizes anzuwenden.

3 **Zum Gebrauch des Wortes „empirisch“:** Das Adjektiv „empirisch“ bei Median, Standardabweichung, Verteilungsfunktion usw. dient zur Unterscheidung solcher aus dem ermittelten Wert (der Realisation) eines n -Tupels von Zufallsvariablen zu berechnenden Kennwerten im Gegensatz zu den entsprechenden Parametern der Wahrscheinlichkeitsverteilung einer Zufallsvariablen in Abschnitt 3.1 bzw. Nr 2.1.3 und Nr 4.3. Diese Unterscheidung kann man durch Hinzufügen des Adjektivs „theoretisch“ zum Parameternamen hervorheben: theoretische Varianz im Gegensatz zur empirischen Varianz. Das Adjektiv „empirisch“ kann weggelassen werden, wenn keine Verwechslungsgefahr mit dem gleichgenannten Parameter besteht

oder durch andere Adjektive oder Zusätze die Verwechslungsgefahr ausgeschlossen wird, z.B. die gemessene Standardabweichung, der Median der Meßreihe usw.

Zu 1.2.10 und 1.2.12

Der Normierungsfaktor $1/(n-1)$ wurde gewählt, damit gilt:

$$E s_X^2 = \text{Var } X_i = \sigma_{X_i}^2 \text{ für } i = 1, \dots, n$$

$$E s_{XY} = \text{Cov}(X_i, Y_i)$$

wenn $X_1, \dots, X_n$ stochastisch unabhängig sind und gleiche Erwartungswerte und Varianzen besitzen bzw. $(X_1, Y_1), \dots, (X_n, Y_n)$ stochastisch unabhängig sind und gleichen Erwartungswert und Kovarianzmatrix besitzen. Sind die in den beiden Fällen genannten Voraussetzungen nicht näherungsweise erfüllt, so werden dem Problem angemessene Normierungsfaktoren gewählt, die jeweils explizit angegeben werden sollten. Um Verwechslungen auszuschließen, empfiehlt es sich, dann ein anderes Zeichen anstelle von s bzw. s^2 zu wählen, siehe insbesondere Nr 1.2.9.1.