Dame-Schule und zeitgenössische Unterhaltungsware. Der Verfasser möchte auf den Versuch, diese Wahl zu rechtfertigen, verzichten, und vielmehr einbekennen, daß Untersuchungen über andere Epochen und Stile noch kaum begonnen wurden. Auf einige Aspekte des Problems hat Heinrich Besseler[25] schon 1938 hingewiesen, so auf die Klavierfeindlichkeit der Innenarchitektur der italienischen Wohnung mit ihren Steinfliesen und Terrazzoböden und auf den Zusammenhang von Mehrchörigkeit und Kirchenarchitektur um 1600.

Methodisch wäre zur Bestimmung der Rolle der Raumakustik noch anzumerken, daß auch hier — den Gedanken Max Webers folgend — zwischen soziologischen Tatbeständen, soziologisch bedingten Tatbeständen und soziologisch relevanten Tatbeständen zu unterscheiden ist, obgleich diese Unterscheidung angesichts der vorliegenden Wechselwirkung nicht immer leicht fällt. So erweist sich die akustische Charakteristik der gotischen Kathedrale deutlich als durch die Architektur und die Gliederung der singenden und hörenden Menschen im Raum bedingt, während die moderne, raum- und elektro-akustisch relativ frei manipulierbare Klangcharakteristik eindeutig als bloß soziologisch relevanter Tatbestand erkennbar ist, der jedoch — wie hier zu zeigen versucht wurde — selbst wieder ein soziologischer Faktor werden kann, denn die aktuelle Wirksamkeit dieses Faktors emanzipiert sich von dessen Genese.

Zuletzt sei noch erwähnt, daß zum Forschungsgebiet Raumakustik und Musiksoziologie nicht nur die Untersuchung der Beziehung von Klangraum und Klangerlebnis gehört, sondern auch die eminent wichtige Frage des Einflusses der Raumakustik auf die tonale Struktur, die praktischen Probleme der raumakustischen Gestaltung neuer Konzertsäle, Theater und Kirchen sowie die Fragen der akustischen Innenarchitektur unserer Wohnräume.

Zur Anwendung quantitativer Methoden in der Tonpsychologie

VON VOLKER RAHLFS, HAMBURG

„Wie so oft, frage ich mich hier, warum denn nicht die Wahrscheinlichkeitsrechnung als Schutz gegen willkürliche Annahmen angerufen wird?“ Dieser Stoßseufzer Handschins[1] führt uns geradewegs zum Kern unseres Problems, nämlich zu der Frage, ob quantitative Methoden in der Ton- und Musikpsychologie überhaupt sinnvoll anzuwenden sind. Von Pythagoras über Kepler bis zu Hans Kayser, immer wieder trachteten feinsinnige Denker danach, Planetensystem und musikalische Harmonie mit Zahlen zu erfassen und möglichst auf die gleichen einfachen Proportionen zurückzuführen. Erhebliche Abweichungen mußten dabei in Kauf genommen werden. Gegen solch spekulatives Denken vor allem wendet sich Handschin, wenn er eine Überprüfung der Hypothesen mit Hilfe der Wahrscheinlichkeitsrechnung empfiehlt. Zwar soll nicht geleugnet werden, daß spekulatives Denken beachtliche Ergebnisse

---

25 Heinrich Besseler: *Musik und Raum* in *Musik und Bild*, Festschrift Max Seiffert zum siebzigsten Geburtstag, 1938, S. 153 und 156.

1 Jacques Handschin, *Der Toncharakter*, Zürich 1948, S. 165.

hervorgebracht hat; wie oft aber soll es eine nüchterne Abschätzung der Tatsachen ersetzen! Es ist kein Zufall, daß gerade die moderne Psychologie in immer stärkerem Maße von der Wahrscheinlichkeitsrechnung und der aus ihr entwickelten Statistik Gebrauch macht. Statistische Überlegungen dringen zunehmend auch in die Tonpsychologie ein. Daher sollte der Musikwissenschaftler, der an tonpsychologischen Fragestellungen interessiert ist, mit den Grundlagen dieser Hilfswissenschaft vertraut gemacht werden. So wendet sich dieser Aufsatz an den Nichtfachmann; mathematische Ableitungen werden zugunsten einer anschaulichen Erläuterung der wichtigsten Begriffe möglichst vermieden.

Die quantitativen Methoden der Psychologie entstammen historisch gesehen zwei Quellen: Einmal der Psychophysik des 19. Jahrhunderts, deren Konzeption vor allem mit den Namen E. H. Weber und G. Th. Fechner verbunden ist[2], zum andern der erst in neuerer Zeit namentlich in Amerika entwickelten Psychometrie. Nahezu alle Phänomene der Wahrnehmung können — wie die amerikanische Forschung gezeigt hat — ohne Rückgriff auf physikalische Reizstrukturen und deren Dimensionen quantifiziert und skaliert werden[3]. Das Hineintragen mathematischer Methoden in die Psychologie bedeutet aber keinesfalls die Wiederaufnahme alter Zahlenmystik; es handelt sich hier um etwas völlig anderes, nämlich um die Ausnutzung der Isomorphierelation zwischen mathematischen und psychologischen Sachverhalten[4]. Die logische Sprache der Mathematik eignet sich wie keine andere zur Beschreibung psychologischer Gesetzmäßigkeiten. Daraus ergibt sich aber zugleich die Einschränkung, daß alle mathematischen Modelle nur Annäherungen sein können.

Der unvorbelastete Leser, dem das Wort Statistik bisher wohl vornehmlich im Zusammenhang mit dem Wirtschaftsleben bedeutungsvoll erschien, mag zunächst wissen wollen, worin der Nutzen besteht, den man sich aus der Anwendung quantitativer Methoden in der Psychologie verspricht. Die präzise Aussage der Statistik bietet unübertroffene Vorteile:

a) Sie erlaubt eine außerordentlich exakte Beschreibung der Sachverhalte;

b) Es können auf diesem Wege allgemeine Schlüsse gezogen werden; auch läßt sich abschätzen, wieviel Vertrauen einer Schlußfolgerung zu schenken ist;

c) Sie ermöglicht Voraussagen;

d) Sie erlaubt, Ursache-Wirkungs-Zusammenhänge in einem Komplex aufzudecken (Faktoren-Analyse).

Die Statistik umfaßt zwei Hauptgebiete: Die deskriptive Statistik und die Stichprobenstatistik.

Deskriptive Statistik

Hierher gehören graphische Darstellungen oder auch Maße, mit denen bestimmte Situationen zahlenmäßig erfaßt werden.

---

[2] Vgl. hierzu u. a. J. C. Flugel, *Probleme und Ergebnisse der Psychologie*, Stuttgart o. J.
[3] W. Witte, *Einführung in die mathematische Behandlung psychologischer Probleme*. In: F. Dorsch, *Psychologisches Wörterbuch*, Hamburg/Bern 1959, S. 397 ff.
[4] Vgl. hierzu W. Witte, a. a. O., S. 405; ebenso J. P. Guilford, *Psychometric Methods*, New York 1954, S. 6—7.

1. Durchschnitt (arithmetisches Mittel).

Mit Hilfe des arithmetischen Mittels lassen sich zwei Gruppen oder Stichproben auf einer gemeinsamen Skala vergleichen. Sein zahlenmäßiger Wert ergibt sich aus der Summe der Meßwerte X_i, geteilt durch ihre Anzahl N:

$$M = \frac{\sum_{i=1}^{N} X_i}{N} \tag{1}$$

Σ ist das mathematische Symbol für eine Summe, wobei der Index i=1 bis N angibt, über welche Meßwerte summiert werden soll.

2. Streuung.

Sehr oft interessiert neben dem Mittelwert auch die Variabilität (Streuung) der Maßzahlen, d. h. die Frage, wie dicht sich die einzelnen Werte um den Mittelwert scharen. Aus Gründen, auf die wir später zurückkommen, empfiehlt sich als Streuungsmaß vor allem die sogenannte Varianz, bzw. ihre Wurzel, die Standardabweichung (SD-Streuung). Hierzu werden die Abweichungen der Meßwerte (X) vom arithmetischen Mittel (M) quadriert; die Summe ihrer Quadrate wird durch (N—1) geteilt.

$$\text{Varianz} = S^2 = \frac{\sum_{i=1}^{N} (X_i - M)^2}{N-1}$$

$$\text{Standardabweichung} = \sqrt{S^2} = S = \sqrt{\frac{\Sigma (X_i - M)^2}{N-1}} \tag{2}$$

3. Korrelation

Zwei Maßreihen der gleichen Gruppe von Individuen oder Beobachtungen weisen mitunter einen Zusammenhang auf, der sich zahlenmäßig durch den Korrelationskoeffizienten darstellen läßt. Sein Wert reicht von +1 (hohe Korrelation) über 0 (keine Korrelation) bis zu —1 (hohe negative Korrelation; sehr enger Zusammenhang in dem Sinne, daß großen X-Werten kleine Y-Werte entsprechen und umgekehrt). Die sogenannte „Produkt-Moment-Korrelation" läßt sich wie folgt berechnen:

$$r_{xy} = \frac{N \Sigma XY - (\Sigma X) \cdot (\Sigma Y)}{\sqrt{[N \cdot \Sigma X^2 - (\Sigma X)^2][N \cdot \Sigma Y^2 - (\Sigma Y)^2]}} \tag{3}$$

X sind die einzelnen Maßzahlen der einen Gruppe, Y die der anderen Gruppe, N ist die Anzahl der verglichenen Maßzahl-Paare (XY).

Graphische Darstellung, Mittelwert, Streuung und Korrelation mögen jetzt an einem Beispiel aus der Tonpsychologie demonstriert werden. Es handelt sich um das von Weber wohl erstmals beschriebene Phänomen der „Tonhöhenverschiebung bei

Intensitätsänderung des Reizes", welches durch Zurmühl[5], Stevens[6] und später durch Kraffczyk[7] eingehend untersucht wurde.

Tabelle der Messungen

Bezugsfrequenz: 400 Hz, Lautstärkeunterschied: 30 dB, Schalldruck des lauteren Tones: ca. 70 DIN-Phon[8].

	1. Stichprobe			2. Stichprobe					
1	2	3	4	5	6	7	8	9	10
Vp	Abweichungen in Hz	Abweichungen in Cents	M_1	Abweichungen in Hz	Abweichungen in Cents	M_2	M_1^2	M_2^2	$M_1\ M_2$
Ra	2,4,3	9, 18, 13	13,3	5, 5, 4	22, 22, 18	20,7	176,89	428,49	275,31
La	7,7,1	31, 31, 5	22,3	-1, 7, 5	- 5, 31, 22	16,0	497,29	256,00	356,80
Da	6,6,7	27, 27, 31	28,3	8, 7, 8	35, 31, 35	33,7	800,89	1135,69	953,71
Krü	5,6,6	22, 27, 27	25,3	5, 5, 5	22, 22, 22	22,0	640,09	484,00	556,60
Kö	5,6,7	22, 27, 31	26,7	8, 9, 11	35, 40, 49	41,3	712,89	1705,69	1102,71
We	4,3,4	18, 13, 18	16,3	1, 9, -2	5, 40, - 9	12,0	265,69	144,00	195,60
Hei	3,6,5	13, 27, 22	20,7	4, 3, 5	18, 13, 22	17,7	428,49	313,29	366,39
He	0,4,8	0, 18, 35	17,7	5, 3, 5	22, 13, 22	19,0	313,29	361,00	336,30
Dae	3,3,2	13, 13, 9	11,7	5, 5, 5	22, 22, 22	22,0	136,89	484,00	257,40
Rei	5,9,7	22, 40, 31	31,0	7, 7, 4	31, 31, 18	26,7	981,00	712,89	827,70
Σ Summe			213,3			231,1	4933,41	6025,05	5228,52

Die Tabelle gibt die in zwei Versuchen[8a] ermittelten Frequenz-Abweichungen an, die erforderlich waren, um bei einem Intensitätsunterschied von 50 dB gleichen Tonhöheneindruck hervorzurufen. Die Versuchspersonen wurden aufgefordert, einen Vergleichston geringer Intensität auf subjektiv gleiche Tonhöhe mit einem 400 Hz-Standardton von höherer Intensität zu bringen. Die Abweichungen konnten mit Hilfe eines stufenweise schaltbaren Meßgenerators[9] auf 1 Hz genau ermittelt werden. Jede Versuchsperson hatte die beiden Töne dreimal zu vergleichen, die Messungen wurden in Spalte 2 eingetragen. Der ganze Versuch wurde nach einer Woche wiederholt (Spalte 5).

---

[5] G. Zurmühl, *Abhängigkeit der Tonhöhenempfindung von der Lautstärke und ihre Beziehung zu Helmholtzschen Resonanztheorie des Hörens.* Zs. f. Sinnesphysiol. 61, 1930, S. 40 ff.

[6] S. S. Stevens, *The relation of pitch to intensity*, J. Acoustic Soc. Am. 6, 1935, S. 150—154.

[7] G. Kraffczyk, *Beobachtungen über Tonhöhenänderungen bei Intensitätsänderungen. Der inogene Frequenzsprung.* Diss. Erlangen 1950.

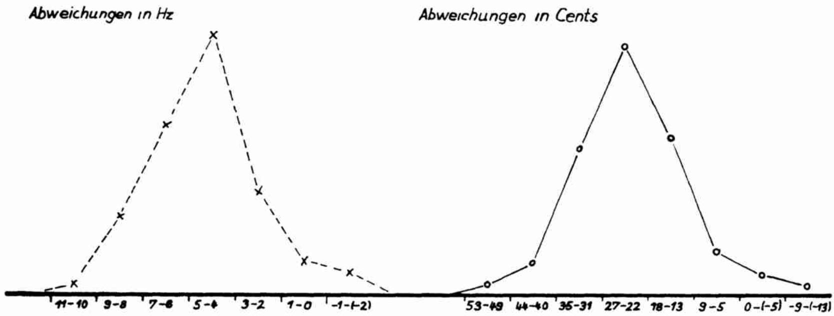
[8] Kontrolliert mit Hilfe eines DIN-Schallpegelmessers Typ EZGN (Rohde & Schwarz).

[8a] Die Versuche wurden im Februar 1962 in einer Übung von Dr. Hanspeter Reinecke mit zehn Studenten und Doktoranden des Musikwissenschaftlichen Instituts der Universität Hamburg durchgeführt.

[9] Dekadischer Meßgenerator, Typ GMG—3 (Wandel & Goltermann).

Zunächst wird man fragen, wie sich die Werte über den Meßbereich verteilen. Die Abweichungen (in Hz) erstrecken sich von +11 bis —2. Aus den Zahlen läßt sich eine Häufigkeitstabelle zusammenstellen, indem die Strecke von +11 bis —2 in 7 Maßzahl-Bereiche (Klassen, Kategorien) aufgeteilt und dann ausgezählt wird, wieviele Fälle auf jede Klasse entfallen. Nehmen wir an, daß die beiden Versuche nicht zu unterschiedlich sind und mischen die Maßzahlen, so ergibt sich folgende Häufigkeitstabelle:

Klasse	11-10	9-8	7-6	5-4	3-2	1-0	(-1)-(-2)
Häufigkeit	1	7	15	23	9	3	2



Figur 1 zeigt die Häufigkeit in den einzelnen Bereichen. Auf der Abszisse sind die Maßzahlen aufgetragen, auf der Ordinate die Häufigkeiten (Fallzahlen). Die Punkte in der Mitte des Bereiches, die die Häufigkeiten darstellen, sind miteinander zu einem Kurvenzug verbunden (Polygonzug).

Im Diagramm erscheinen auf der Abszisse die Klassen; die Ordinate gibt die jeweilige Häufigkeit der Fälle an. Es entsteht eine rechtsgipflige schiefe Kurve.
Von der Psychometrie her ist bekannt, daß sich beim Messen von psychischen Erscheinungen mit Hilfe physikalischer Maße (in unserem Falle Hz) häufig derart schiefe Verteilungsformen ergeben. Durch geeignete Transformation der Werte kann man diese Schiefe beseitigen.
Das Tonhöhenphänomen ist als eine etwa logarithmische Entsprechung von „innen und außen" bekannt. Die Maßzahlen werden daher nach der Gleichung
$X_{neu} = \log X_{alt}$ umgewandelt[10]; für die Tonhöhe empfiehlt sich vor allem die Cents-Rechnung[11].
Die in Cents umgerechneten Werte haben folgende Häufigkeitsverteilung:

Klasse	53-49	44-40	35-31	27-22	18-13	9-5	0-(-5)	(-9)-(-13)
Häufigkeit	1	3	13	22	14	4	2	1

Diese Centswerte, in ein Diagramm übertragen, zeigen nunmehr eine recht gute Annäherung an eine ideale Kurve, die sich symmetrisch um den Mittelwert erstreckt

10 Erich Mittenecker, *Planung und statistische Auswertung von Experimenten*, 2. Aufl., Wien 1958, S. 71.
11 Sie erlaubt den anschaulichen Vergleich beliebiger Distanzen mit dem als Grundmaß gesetzten temperierten Halbtonschritt = 100 Cents (Heinrich Husmann, *Fünf- und siebenstellige Centstafeln zur Berechnung musikalischer Intervalle*, Ethnomusicologica, Vol. II, Leiden 1951).

(Figur 1b). Eine derartige charakteristische Verteilung, die der sogenannten Normalkurve (Glockenkurve) entspricht, kommt immer dann zustande, wenn in die meßbaren Merkmale zahlreiche voneinander unabhängige Faktoren eingehen. Wir nehmen jetzt die Normalverteilung der Centswerte als gegeben an[12] und beginnen mit der Berechnung der sogenannten Statistiken.

Mittelwert

Zunächst wird der Mittelwert der Betrachtungen jeder Versuchsperson nach Formel (1) berechnet und als „wahrer" Wert in Spalte 4 und 7 der Tabelle eingetragen. Da aber vor allem die Ergebnisse der Gruppe interessieren, wird aus den 10 Mittelwerten jetzt der Gruppen-Mittelwert gebildet. Hierzu wird die Summe aller 10 Werte durch 10 dividiert. Bei der 1. Stichprobe beträgt der Durchschnitt 21,3 Cents (Spalte vier der Tabelle), bei der 2. Stichprobe 23,1 Cents.

Streuung

Die Streuung des ersten Mittelwertes ergibt sich aus den Werten der Spalte vier, die in die Formel (2) oder in die mathematische völlig gleichwertige Formel

$$s = \sqrt{\frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N-1}} \tag{4}$$

eingesetzt werden.

ΣX^2 finden wir in Spalte 8 in der Summenzeile, ΣX in Spalte 4 in der Summenzeile. N ist die Anzahl der summierten Werte, in diesem Falle gleich 10.

Wir erhalten also:

$$s = \sqrt{\frac{4933{,}4 - \frac{(213{,}3)^2}{10}}{9}} = 6{,}53$$

für die zweite Stichprobe:

$$s = \sqrt{\frac{6025{,}05 - \frac{(231{,}3)^2}{10}}{9}} = 7{,}60$$

Der Wert „s" hat neben der mathematischen auch eine rein anschauliche und praktische Bedeutung: Für die Normalverteilung (Gaußsche Glockenkurve) gilt, daß etwa $^2/_3$ aller Fälle im Bereich von s beiderseits des Mittelwerts liegen. Das bedeutet: Im ersten Versuch liegen etwa $^2/_3$ aller Beobachtungen im Bereich 21,3 $\pm$ 6,53 Cents, also zwischen 14,77 und 27,83 Cents; im 2. Versuch entsprechend 23,1 $\pm$ 7,60 Cents.

---

12 Eine genaue statistische Überprüfung dieser Annahme ist mit Hilfe des sog. Chi-Quatrats möglich. Vgl. hierzu: Peter R. Hofstätter, *Einführung in die quantitativen Methoden der Psychologie*, München 1953, ebenso J. P. Guilford, *Fundamental Statistics in Psychology and Education*, New York, 1956.

Da es sich bei beiden Versuchen um den gleichen Personenkreis handelt, ist ein Zusammenhang der Werte zu erwarten. Die Korrelation beider Maßreihen ergibt sich durch Einsetzen der entsprechenden Tabellen-Werte in Formel (3):

$$r_{M_1M_2} = \frac{10 \cdot 5228{,}52 - (213{,}3 \cdot 231{,}1)}{\sqrt{[10 \cdot 4933{,}41 - 213{,}3^2]\,[10 \cdot 6025{,}05 - 231{,}1^2]}} = 0{,}58$$

Der Korrelationskoeffizient deutet in diesem Fall auf einen gewissen Zusammenhang hin.

Stichprobenstatistik

Die Stichprobenstatistik ermöglicht vor allem zwei Operationen:

a) Schluß von Stichprobenstatistiken auf die Parameter der Grundgesamtheit (Population), der die Stichprobe entnommen worden ist. Jeder noch so umfangreiche Versuch ist immer nur eine Stichprobe, da die Gesamtheit nur selten erfaßt werden kann.

b) Prüfung eines Unterschiedes auf seine Verläßlichkeit, die über Verwerfung oder Annahme einer Hypothese entscheidet. In einer empirischen Wissenschaft läßt sich eine solche Entscheidung nur mit größerer oder geringerer Sicherheit (Verläßlichkeit) treffen. Man rechnet mit einem Verläßlichkeitsniveau (α), meist mit dem sogenannten 5-Prozent-Niveau ($\alpha = 0{,}05$) oder mit dem 1-Prozent-Niveau ($\alpha = 0{,}01$). Das besagt: Eine Abweichung in weniger als 5 % (bzw. 1 %) der Fälle wird durch das bloße Spiel des Zufalls entstanden sein; man bezeichnet sie als „nicht signifikant".

Bei der in der Psychologie oft vorkommenden Normalverteilung von Daten steht die Häufigkeit in einer festen Beziehung zur entsprechenden Maßzahl:

$$Y = \frac{N}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{X-M}{\sigma}\right)^2} \tag{5}$$

Hierbei gilt: Y = Häufigkeit, N = Anzahl der Messungen, σ = Standardabweichung[13], $\pi = 3{,}14$, $e = 2{,}718$, X = Maßzahl, M = Mittelwert.

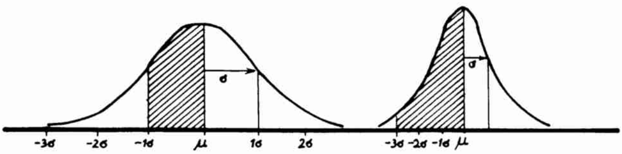
Durch Einsetzen der Zahlen ergibt sich:

$$Y = \frac{N}{2{,}5066\,\sigma}\; 2{,}718^{-\frac{1}{2}\left(\frac{X-M}{\sigma}\right)^2}$$

N, M und σ (bzw. s) sind bei jeder Stichprobe bekannt, es kann also jede beliebige Maßzahl X eingesetzt und ihre Häufigkeit ausgerechnet werden. Die Zuordnung von Maßzahl und Häufigkeit läßt sich graphisch übersichtlich darstellen (Fig. 2)

---

[13] Üblicherweise bezeichnet man in der neueren Literatur die Werte der Population im Gegensatz zu Stichprobenwerten mit griechischen Buchstaben.



Figur 2. Die Gaußsche Normalverteilung. Sie ist durch μ (M) und σ hinreichend definiert. Die Ordinate bei der Maßzahl 1σ gibt die Häufigkeit der Fälle für diesen Wert an, die schraffierte Fläche die Häufigkeit von dieser Maßzahl bis zum Mittelwert.

Der Mittelwert ist zugleich der häufigste Wert, mit zunehmender Entfernung von der Symmetrieachse = M wird die Häufigkeit immer geringer. Sie läßt sich berechnen, indem man die Abweichung der Maßzahl von M in Einheiten von σ mißt. Der Quotient aus (X—M) und σ bildet den sogenannten z-Wert, von dem aus auf die Häufigkeit geschlossen werden kann[14]. Als Beispiel diene ein Meßwert der ersten Stichprobe: Der Mittelwert betrug 21,3 Cents, die Streuung 6,53 Cents. Die Häufigkeit der Klasse um den Meßwert 31 ergibt sich aus

$$z = \frac{X - M}{\sigma} = \frac{31 - 21{,}3}{6{,}53} = 1{,}49$$

nach der z-Tabelle also 0,13, auf N = 60 Fälle bezogen ergeben sich 18 Fälle (Klassenbreite = 9 Cents).

In der Praxis interessiert gewöhnlich die Frage, wie viele Fälle zwischen einem bestimmten Meßwert und dem Mittelwert liegen, in unserem Beispiel also etwa zwischen 31,0 und 21,3 Cents. Diese Häufigkeit ist durch die Fläche unter dem Kurventeil M bis z gegeben[15]. Dem z-Wert unseres Beispiels von 1,49 entspricht laut z-Tabelle (Flächenspalte) 43,19 %. Dieser Wert muß verdoppelt werden, wenn die Fälle beiderseits des Mittelwerts interessieren. Im Bereich 21,3 ± 9,7 Cents, d. h. zwischen 11,6 und 31 Cents, liegen 86,38 % aller Beobachtungen.

In Übereinstimmung mit dem Verläßlichkeitsniveau, von dem schon die Rede war, sind die Bereiche, die 95 % bzw. 99 % aller Fälle umfassen, besonders bedeutsam. Der restliche Teil wird dagegen willkürlich vernachlässigt. Unter Berücksichtigung der z-Beziehung ergibt sich der gewünschte Bereich:

$$\begin{aligned} M &\pm 1{,}96 \cdot s \text{ für } 95\,\%\text{ige Sicherheit} = 5\,\%\text{-Niveau;} \\ M &\pm 2{,}58 \cdot s \text{ für } 99\,\%\text{ige Sicherheit} = 1\,\%\text{-Niveau.} \end{aligned} \tag{6}$$

Hiernach befinden sich alle Werte unseres Beispiels mit 95 %iger Sicherheit im Bereich von 21,3 ± 12,80 Cents, also zwischen 8,5 Cents und 34,1 Cents, mit 99 %iger Sicherheit dagegen zwischen 4,45 Cents und 38,15 Cents. Damit zeigt sich, daß die Sicherheit mit der Breite des Bereichs zunimmt.

Was zunächst wie eine methodische Übung aussah, ist für den Schluß von Stichprobenwerten auf die Parameter der Population als Präzisionsbestimmung unerläßlich.

---

14 Für den praktischen Gebrauch können die Häufigkeiten direkt aus der z-Tabelle entnommen werden.

15 Sie läßt sich in einer eigenen Spalte der z-Tabelle ablesen. Die Gesamtfläche unter der Kurve wird im allgemeinen gleich 100 % gesetzt; somit ist die Tabelle für jede Anzahl der Messungen N geeignet.

Der Bereich, in den der Mittelwert der Population fällt, wird wiederum mit Hilfe des z-Wertes errechnet, da auch die Mittelwerte aller beliebigen Stichproben normal verteilt sind. Bei der Schätzung des Populationsmittelwertes gilt es, den Schätzfehler zu berücksichtigen: Dieser ist der Streuung der Population (σ) direkt und der Wurzel der Stichprobengröße (N) umgekehrt proportional. Der Standardfehler des Mittelwertes ist also:

$$s_M = \frac{\sigma}{\sqrt{N}} \tag{7}$$

Da die Streuung der Population (σ) selten bekannt ist, wird sie in der Praxis durch die Stichprobenstreuung (s) ersetzt:

$$s_M = \frac{s}{\sqrt{N}} \tag{7.1}$$

Der Mittelwert sämtlicher Stichproben aus der Population, der auch zugleich Mittelwert der gesamten Population ist, liegt im Bereich:

$$M \pm 1{,}96\, s_M \text{ für 95 \%ige Sicherheit} = 5\,\%\text{-Niveau.} \tag{8}$$

In unserem Beispiel betragen M = 21,3 Cents und s = 6,53 Cents; daher liegt der „wahre" Wert mit großer Wahrscheinlichkeit im Bereich 21,3 ± 3,92 Cents, mithin zwischen 17,38 und 25,22 Cents.

Mindestens ebenso wichtig wie die Abschätzung des Vertrauensbereiches ist die Überprüfung der Signifikanz von Differenzen, d. h. die Entscheidung der Frage, ob ein beobachteter Unterschied bedeutsam (signifikant) oder nur auf das Spiel des Zufalls zurückzuführen ist. Wollen wir etwa wissen, ob der Mittelwertsunterschied zwischen unseren beiden Stichproben signifikant ist, so müssen wir zunächst prüfen, ob sie sich hinsichtlich ihrer Streuung nicht zu sehr voneinander unterscheiden (Prüfung auf Homogenität der Varianz). Auch hier kann ein zufälliger Unterschied vorhanden sein. Um das zu ermitteln, wird das Verhältnis F der beiden Varianzen gebildet:

$$F = \frac{s_1^2}{s_2^2} = \frac{\text{größere Varianz}}{\text{kleinere Varianz}} = \frac{57{,}76}{42{,}64} = 1{,}35 \tag{9}$$

Dieses F liegt noch unterhalb der bei dem 1 %-Niveau gerade noch zulässigen Größe von 2,97 (F-Tabelle); die Differenz der Streuungen kann daher als zufällig gelten. Die beiden Mittelwerte lassen sich mit Hilfe des sogenannten t-Tests vergleichen. Hierzu setzen wir ihre Differenz in das Verhältnis zum Standardfehler der Differenz (Unsicherheit der Differenz). Dieser ergibt sich nach einem Additionstheorem aus:

$$s_{M_1-M_2} = \sqrt{s_{M_1}^2 + s_{M_2}^2} \tag{10}$$

bei korrelierten Stichproben (etwa bei den gleichen Versuchspersonen) aus

$$s_{M_1-M_2} = \sqrt{s_{M_1}^2 + s_{M_2}^2 - 2 r_{M_1 M_2} \cdot s_{M_1} \cdot s_{M_2}} \qquad (10.\ 1)$$

Die Differenz der Mittelwerte im Verhältnis zu ihrer Unsicherheit ergibt den Kritischen Bruch:

$$CR = \frac{M_1 - M_2}{s_{M_1-M_2}} \qquad (11)$$

dessen Wert wieder mit dem gerade noch zulässigen z-Wert verglichen werden kann. Bei kleinen Stichproben ($N < 100$) sind die Daten gewöhnlich nur angenähert normal verteilt. Man benutzt in diesem Fall an Stelle von CR den sogenannten t-Wert, der die Größe der Stichprobe berücksichtigt:

$$t = \frac{M_1 - M_2}{s_{M_1-M_2}} = \frac{M_1 - M_2}{\sqrt{s_{M_1}^2 + s_{M_2}^2}} \qquad (12)$$

In unserem Beispiel benutzen wir wegen des niedrigen N und der Korrelation:

$$t = \frac{M_1 - M_2}{\sqrt{s_{M_1}^2 + s_{M_2}^2 - 2 r_{M_1 M_2} s_{M_1} s_{M_2}}} \qquad (12.\ 1)$$

Durch Einsetzen ergibt sich:

$$t = \frac{23{,}1 - 21{,}3}{\sqrt{7{,}60^2 + 6{,}53^2 - 2 \cdot 0{,}58 \cdot 7{,}60 \cdot 6{,}53}} = 0{,}275$$

Die t-Tabelle liefert unter Berücksichtigung der Stichprobengröße auf dem 1%-Niveau einen Wert von 2,878. Da unser errechneter Wert aber erheblich unter dem angegebenen liegt, behaupten wir, daß der Unterschied der Mittelwerte nur zufällig entstanden ist. Die Nullhypothese, die das Fehlen eines Unterschiedes annimmt ($M_1 - M_2 = 0$), darf also, wie der Statistiker sagt, auf dem 1%-Niveau als gesichert gelten.

An diesem Beispiel statistischer Behandlung eines begrenzten tonpsychologischen Versuchs sollte gezeigt werden, welche Hilfsmittel die Statistik überhaupt zur Verfügung stellt. Neben der reinen Beschreibung statistischer Operationen konnten zugleich zwei grundsätzliche Probleme eines jeden tonpsychologischen Versuchs in Angriff genommen werden: Die Festlegung des „wahren" Wertes und die Prüfung der Signifikanz. Einerseits ließ sich bestimmen, in welchem Bereich der Mittelwert sämtlicher denkbarer Fälle der Tonhöhenabweichungen unter gleichen Bedingungen mit großer Wahrscheinlichkeit zu suchen ist. Andererseits konnte festgestellt werden, daß trotz der abweichenden Ergebnisse zweier zu verschiedenen Zeiten durchgeführter Versuchsreihen statistisch kein Unterschied besteht, da die Differenz un-

terhalb der durch die Normalverteilung gegebenen Grenzen liegt und deshalb als zufällig anzusehen ist.
Hier konnte es natürlich zunächst nur um die theoretischen Grundgedanken gehen; es ist klar, daß sich ein derart umfangreiches Gebiet nicht auf wenigen Seiten erschöpfend behandeln läßt. Völlig zu kurz kommen mußte das Planen von Versuchen, eine eigene Methodik, die im angelsächsischen Schrifttum unter dem Namen „experimental design" großen Platz einnimmt[16].
Schließlich sei noch vermerkt, daß sich die moderne Statistik nicht auf die dargestellten Möglichkeiten beschränkt, sondern für die Besonderheiten des psychologischen Experiments völlig neue Methoden entwickelt hat. Sie umfassen in der Hauptsache drei Gebiete:

1. Nichtparametrische Methoden: statistische Prüfverfahren, die nicht von einer besonderen Art der Verteilung (etwa der Normalverteilung) abhängen[17].

2. Varianzanalyse: Sie dient der Ermittlung der Verläßlichkeit von Unterschieden zwischen kleineren Gruppen.

3. Faktorenanalyse: Ein rechnerisches Verfahren, mit dem aus einer größeren Zahl miteinander korrelierter Variabler eine kleinere Zahl voneinander unabhängiger Variabler (Faktoren) gewonnen wird[18].

Die subtilen Methoden der Varianz- und Faktorenanalyse haben auch in der Tonpsychologie ihre Bewährungsprobe bereits abgelegt[19]. Mit ihrer Hilfe können komplexe Zusammenhänge von Toneigenschaften nunmehr rein psychologisch, d. h. ohne Rückgriff auf Dimensionen der physikalischen Welt, exakt erfaßt werden. Damit entfällt aber die Notwendigkeit der Orientierung an akustischen oder physiologischen Gesetzen, die bisher als quasi objektiver Hintergrund aller Forschung angesehen wurden, jedoch immer wieder den Blick ablenkten von der Eigengesetzlichkeit des musikalischen Erlebens als eines vornehmlich psychologischen Problems[20].

---

16 Vgl. u. a. A. L. Edwards, *Experimental design in psychological research*, 1950.
17 G. A. Lienert, *Die statistische Beurteilung von Gruppenunterschieden durch sogenannte verteilungsfreie Prüfverfahren*, Psychol. Beitr., 1957, S. 38—79; Siegel, *Nonparametric Statistics for the Behavioral Sciences*, New York 1956.
18 P. R. Hofstätter, *Über Faktorenanalyse*, Arch. f. d. ges. Psychol. 1938, 100, S. 223—279; ebenso G. A. Lienert, *Prinzip und Methode der multiplen Faktoren-Analyse, demonstriert an einem Beispiel*, Biometr. Zeitschr., 1959, 1. Thurstone, *Multiple Factor Analysis*, Chicago 1947.
19 Vgl. u. a.: J. E. Karlin, *A factorial study of auditory function*, Psychometrika 7, 1942, S. 251—279; ebenso J. Korboot und J. A. Keats, A *Multidimensional Scaling of Sounds*, Australian Journal of Psychology 11, 1959, S. 62—69.
20 H.-P. Reinecke, *Stellung und Grenzen akustischer Forschung innerhalb der systematischen Musikwissenschaft*, Acta mus. 31, 1959, S. 80—86; H.-P. Reinecke, *Über die Eigengesetzlichkeit des musikalischen Hörens*, Musikalische Zeitfragen, hrsg. v. Walter Wiora, Heft 10, Kassel 1962.