

**Zeitschrift:** Elemente der Mathematik  
**Band:** 5 (1950)  
**Heft:** 3

**Artikel:** Über die Entwicklung des zentralen Grenzwertsatzes der Wahrscheinlichkeitsrechnung  
**Autor:** Saxer, Walter  
**DOI:** <https://doi.org/10.5169/seals-14906>

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

**Download PDF:** 15.10.2024

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

une ponctuelle (2) projective au même faisceau, les ponctuelles (1) et (2) sont projectives; or, de plus, elles sont perspectives: Si  $C$  vient en  $A$ ,  $D$  est à l'infini de  $t$  et son correspondant dans (2) également. Par conséquent, toutes les droites  $DP$  se coupent en un même point. Ce point devant se trouver sur  $TA$  comme on le voit en prenant  $C$  en  $T$ , c'est le point  $R$ . Le lieu du point  $P$  est donc le cercle de diamètre  $TR$ .

Quels sont les points d'intersection de ce cercle avec la conique? En un tel point  $E$ , les points  $P$  et  $C$  coïncident, étant alignés sur  $T$ .  $TE$  et  $ER$  sont perpendiculaires:  $RE$  et  $EB$  devant se couper sur  $t$ , ces deux droites coïncident.  $E$  doit donc se trouver à l'intersection du cercle de diamètre  $TB$  avec la conique. Or trois des points d'intersection sont  $T$ ,  $B$  et  $A$ ; il n'y a donc qu'un seul point  $E$  qui est aussi le point d'intersection (différent de  $T$ ) de la conique, du cercle de diamètre  $TR$  et du cercle de diamètre  $TB$ . Comme le cercle de diamètre  $TR$  ne coupe la conique qu'en un seul point en dehors de  $T$ , c'est le cercle de courbure à la conique en  $T$ , ce qui démontre notre théorème.

Remarques.

a) Si  $T$  est le sommet d'une conique, les points  $A$  et  $B$  coïncident avec le sommet opposé et le théorème reste valable.

b) Si de plus la conique est une parabole, on a le problème n° 67 proposé par M. E. ROTHMUND.

c) Si  $B$  est un point quelconque de la conique,  $T, C, D$  et  $P$  conservant leur signification, le lieu du point  $P$  est alors une cubique de genre zéro ayant un point double en  $T$ .

J.-P. SYDLER, Zurich.

## Über die Entwicklung des zentralen Grenzwertsatzes der Wahrscheinlichkeitsrechnung<sup>1)</sup>

Die Wahrscheinlichkeitsrechnung, einschließlich der mathematischen Statistik, ist eines jener Gebiete der Mathematik, die sich im 20. Jahrhundert und ganz besonders in den letzten zwei Jahrzehnten außerordentlich rasch entwickelt haben und auch bezüglich der Anwendungen in den exakten Naturwissenschaften und der Technik heute zusammen mit der Analysis an vorderster Stelle stehen. Bei Beginn unseres Jahrhunderts spielte die Wahrscheinlichkeitsrechnung in der Mathematik noch eine recht bescheidene Rolle. Sie bestand aus einer Reihe interessanter Einzelresultate, die auf Fragen der Fehlerrechnung, Versicherung und — Spiele angewendet wurden. Von einer eigentlichen Theorie und einem axiomatischen Aufbau konnte man noch nicht sprechen.

Heute nimmt die Wahrscheinlichkeitsrechnung eine bedeutende Stellung in der Mathematik und überhaupt in der Wissenschaft ein. Dank der Schaffung der Lebesgueschen Maß- und Integrationstheorie wurde zunächst ein begrifflich einwandfreier axiomatischer Aufbau der Wahrscheinlichkeitsrechnung möglich. Es war nicht mehr nötig, wie früher zwischen arithmetischen und geometrischen Wahrscheinlichkeiten zu unterscheiden. Mit den sogenannten charakteristischen Funktionen, die mit jeder

<sup>1)</sup> Nach einem Vortrag im Mathematischen Kolloquium Winterthur.

Wahrscheinlichkeitsverteilung verbunden sind, wurde ein sehr zweckmäßiges analytisches Instrument geschaffen, mit dem sich dank der Kenntnisse aus der Fourier-Integraltransformation wichtige neue mathematische Sätze gewinnen ließen.

Neben dem theoretischen Ausbau der Wahrscheinlichkeitsrechnung erfolgte gleichzeitig eine überaus große Ausweitung ihrer Anwendungsmöglichkeiten auf die verschiedensten Gebiete der Naturwissenschaften und der Technik. Die Wahrscheinlichkeitsrechnung kann als eine mathematische Theorie der Massenerscheinungen bezeichnet werden. Solche kommen in den verschiedensten Gebieten, wie statistische Mechanik, Molekülbau, Diffusionserscheinungen, Kolloidchemie, Biologie (wie Trefprinzip in der Medizin), Massenfabrikation und ihre statistische Kontrolle, Fernmeldetechnik, Theorie statistischer Reihen, wie z. B. Zeitreihen usw., vor. Aus diesem Grunde hat sich die mathematische Statistik in den letzten Jahren auffallend entwickelt; sie wurde im Interesse der wissenschaftlichen Aufrüstung in den angelsächsischen Ländern besonders ausgiebig gefördert (vgl. z. B. die Memoiren WINSTON CHURCHILLS) und ist heute beinahe eine selbständige Disziplin geworden.

Schließlich sei darauf hingewiesen, daß die Wahrscheinlichkeitsrechnung auch in der Philosophie der Wissenschaften eine bedeutende Stellung einnimmt.

Im folgenden soll versucht werden, an Hand der Entwicklung des zentralen Grenzwertsatzes der Wahrscheinlichkeitsrechnung ein Bild von ihrem heutigen Stand zu geben. Dieser wichtigste Satz hat seinen Ursprung in der Binomialverteilung von NEWTON-BERNOULLI. Bezeichnet  $p$  die Wahrscheinlichkeit des Eintreffens eines bestimmten Ereignisses (Ziehung eines bestimmten Loses aus einer Urne), so beträgt die Wahrscheinlichkeit  $w_n(v)$ , bei  $n$ -maliger Wiederholung dieser Ziehung unter gleichen Voraussetzungen  $v$ -mal dieses Los zu ziehen:

$$w_n(v) = \binom{n}{v} p^v (1-p)^{n-v}.$$

$v$  kann in diesem Falle die Werte  $0, 1, \dots, n$  annehmen. Die Wahrscheinlichkeitsverteilung von  $v$  ist durch die obige Formel geregelt. Die sogenannte mathematische Erwartung  $E(v)$  von  $v$ , der Mittelwert dieser stochastischen Variablen (Variable, deren Werte mit einer gewissen Wahrscheinlichkeit angenommen werden) beträgt

$$E(v) = \sum_{v=0}^n v w_n(v) = n p.$$

Die Streuung  $\text{Str}(v)$  dieser Variablen um ihren Mittelwert herum beträgt:

$$\text{Str}(v) = \sum_{v=0}^n w_n(v) [v - n p]^2 = n p q.$$

Schon DE MOIVRE betrachtete im Jahre 1733 an Stelle der Variablen  $v$  die standardisierte Variable  $x$ :

$$x = \frac{v - n p}{\sqrt{n p q}}$$

(Mittelwert dieser Variablen 0, ihre Streuung 1)

und bewies durch Grenzübergang ( $n \rightarrow \infty$ ) den folgenden Satz: Die vorhin erwähnte

Binomialverteilung geht durch diese Transformation – verbunden mit dem Grenzübergang – zu einer stetigen Wahrscheinlichkeitsverteilung über. Das heißt:  $x$  kann alle Werte von  $-\infty$  bis  $+\infty$  mit einer gewissen Wahrscheinlichkeit annehmen. Die Wahrscheinlichkeit dafür, daß  $x$  die Werte zwischen  $-\infty$  und  $x$  annehme, beträgt:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt.$$

Die so definierte Wahrscheinlichkeitsverteilung ist bekanntlich die Normalverteilung oder Gaußsche Verteilung. Sie ist die wichtigste Verteilung und spielt in der ganzen Wahrscheinlichkeitsrechnung samt Anwendungen eine zentrale Rolle. Aus diesem Grunde wurde die obige Funktion  $\Phi(x)$  samt ihren Ableitungen tabellarisiert.

Bereits das Ergebnis von DE MOIVRE kann in der folgenden Weise interpretiert werden: Statt  $v$ -mal das Los zu ziehen, kann jede Ziehung durch eine stochastische Variable dargestellt werden, die lediglich die Werte 1 und 0 annimmt (Treffer oder kein Treffer).  $n$  Ziehungen bedeuten die Summenbildung von stochastischen gleichen Variablen; das de-Moivresche Resultat gibt die Verteilung dieser Summe, nachdem noch eine gewisse, zweckmäßige Transformation der Variablen vorgenommen wurde.

Es soll nunmehr beschrieben werden, in welcher Form dieser Grenzwertsatz in der modernen Wahrscheinlichkeitsrechnung formuliert und welches seine wahre Bedeutung für die Anwendungen ist.

Mit  $\xi$  werde eine stetige stochastische Variable bezeichnet, die alle Werte zwischen  $-\infty$  und  $+\infty$  annehmen dürfe. Ihre Verteilung werde durch ihre Verteilungsfunktion  $F(x)$  definiert, d. h. es sei die Wahrscheinlichkeit  $W(\xi, -\infty < \xi \leq x)$ , wobei  $F(x)$  eine wachsende Funktion ist.

$$F(-\infty) = 0, \quad F(+\infty) = 1.$$

Sofern  $F(x)$  stetig und eine erste Ableitung  $f(x)$  besitzt, bedeutet  $f(x)$  die sogenannte Wahrscheinlichkeitsdichte. Die Wahrscheinlichkeit, daß  $\xi$  zwischen den Werten  $x$  und  $x + dx$  liegt, beträgt  $f(x) dx$ . Unter der mathematischen Erwartung  $E(\xi)$  von  $\xi$  versteht man den Ausdruck:

$$E(\xi) = \int_{-\infty}^{+\infty} x f(x) dx,$$

und unter der Streuung:

$$D^2(\xi) = \int_{-\infty}^{+\infty} f(x) [x - E(\xi)]^2 dx.$$

Allgemein versteht man unter den Momenten einer Verteilung bezüglich ihres Mittelwerts die Größen

$$E[\xi - E(\xi)]^v = \int_{-\infty}^{+\infty} f(x) [x - E(\xi)]^v dx.$$

Mathematische Erwartung und Streuung sind mit dem ersten und zweiten Moment der Verteilung identisch.

Unter der charakteristischen Funktion  $\varphi(t)$  einer Wahrscheinlichkeitsverteilung versteht man den Erwartungswert der Größe  $e^{it\xi}$ , d. h.

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx.$$

Mit Hilfe fundamentaler Sätze aus der Theorie der Fourier-Integraltransformation kann man zeigen, daß unter sehr allgemeinen Voraussetzungen Eineindeutigkeit zwischen einer Wahrscheinlichkeitsverteilung und ihrer charakteristischen Funktion besteht. Für die Anwendungen der charakteristischen Funktion ist nun die Bemerkung wichtig, daß gewisse Operationen bei stochastischen Variablen — wie ihre Überlagerung — sich bei der charakteristischen Funktion in einfachen, übersichtlichen Operationen auswirken. Es besteht hier eine gewisse Analogie zur Laplace-Transformation. Aus diesem Grunde und der Eineindeutigkeit zwischen der Verteilungsfunktion und ihrer charakteristischen Funktion wird dann der Einfluß von Rechnungsoperationen zuerst bei der charakteristischen Funktion überprüft.

Um den zentralen Grenzwertsatz nach dem heutigen Stand der mathematischen Erkenntnis zu formulieren, muß man zunächst den Begriff der Normalverteilung durch die folgende Funktion  $\Phi(m, \sigma, x)$  allgemeiner formulieren:

$$\Phi(m, \sigma, x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(\xi - m)^2}{2\sigma^2}} d\xi.$$

Diese Normalverteilung besitzt den Mittelwert  $m$  sowie die Streuung  $\sigma$  und geht mit Hilfe einfacher Transformationen aus  $\Phi(x)$  hervor.

Schon LAPLACE hat vermutet, daß der folgende Satz richtig sei: Die Summe von  $n$  stochastischen unabhängigen Variablen  $\xi_1 + \dots + \xi_n$  mit den Mittelwerten  $m_1, m_2, \dots, m_n$  und den Streuungen  $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$  konvergiert unter sehr allgemeinen Voraussetzungen über die Verteilungen der  $\xi_i$  gegen eine Normalverteilung mit dem Mittelwert  $m_1 + \dots + m_n = m$  und der Streuung

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2.$$

Dieser zum erstenmal vom russischen Mathematiker LIAPOUNOFF exakt bewiesene Satz verlangt nicht mehr wie bei DE MOIVRE, daß die  $\xi_i$  gleich, und zwar gemäß der Newtonschen Formel verteilt seien. Die Frage, welche Voraussetzungen über die Verteilung der  $\xi_i$  für die Gültigkeit des zentralen Grenzwertsatzes gefordert werden müssen, wurde in den letzten Jahren sehr genau untersucht. Schließlich gelang es dem heute in den USA lebenden deutschen Mathematiker FELLER, diesen Grenzwertsatz als reinen Konvergenzsatz der Analysis zu betrachten und die folgende Frage abschließend zu beantworten:

Welche Bedingungen müssen die Zahlfolgen  $a_1, a_2, \dots$  und  $b_1, b_2, \dots$  sowie die Verteilungsfunktionen  $F_i(\xi_i)$  der stochastischen Variablen  $\xi_i$  erfüllen, damit die Verteilungsfunktion von  $u_n$ :

$$u_n = \frac{\xi_1 + \dots + \xi_n - b_n}{a_n},$$

gegen  $\Phi(x)$  konvergiert?

Der zentrale Grenzwertsatz ist für die Anwendungen von fundamentaler Bedeutung, und zwar aus zwei Gründen: Viele stochastische Variablen der Natur lassen sich als Überlagerungsergebnis vieler Einflüsse, als Summe vieler stochastischer Variablen betrachten. Dies ist die Erklärung dafür, daß sehr viele Größen normal verteilt sind. Aus diesem Grunde hat schon GAUSS Fehler gemessener oder beobachteter Größen als normal verteilt angenommen. Beispielsweise hat gerade einer meiner Doktoranden festgestellt, daß die durchschnittliche bei schweizerischen Pegelstationen gemessene jährliche Wassermenge für die wichtigsten schweizerischen Flüsse zeitlich normal verteilt ist, daß dies aber für die Halbjahresmenge aus bestimmten Gründen nicht ohne weiteres zutreffend ist. Im übrigen soll man sich trotz der großen Bedeutung der Normalverteilung davor hüten, stets Normalverteilungen bei stochastischen Variablen der Natur anzunehmen. Die moderne mathematische Statistik hat gute Methoden ausgearbeitet, um den Typus einer aus der Natur gegebenen Verteilung möglichst sicher festzustellen.

Schon DANIEL BERNOULLI hat die Frage nach einem universellen Fehlergesetz aufgeworfen. Er vermutete mit Recht, daß doch ganz bestimmte mathematische Gründe für die Kongruenz solcher Fehlerverteilungen mit einer Normalverteilung vorhanden sein müssen. Diese mathematische Erklärung wurde in den letzten Jahrzehnten gefunden und liegt im wesentlichen darin, daß die Normalverteilung das einzige Fehlergesetz, die einzige Verteilung mit endlicher Streuung ist, die bei linearer Zusammensetzung der Fehler erhalten bleibt. Darunter verstehen wir folgendes:

Jede Funktion

$$\varphi(x) \geq 0 \quad \text{und} \quad \int_{-\infty}^{+\infty} \varphi(x) dx = 1$$

werde als ein Fehlergesetz bezeichnet. Man bemerkt sofort, daß mit  $\varphi(x)$  auch  $1/a \cdot \varphi(x/a)$  mit  $a > 0$  ebenfalls ein Fehlergesetz darstellt. Etwas versteckter ist die Aussage, daß auch das Faltungsintegral

$$\int_{-\infty}^{+\infty} \varphi(u) \psi(x-u) du$$

ebenfalls ein Fehlergesetz ist, wenn dies für  $\varphi(x)$  und  $\psi(x)$  zutrifft. (Es handelt sich beim Faltungsintegral einfach um die Überlagerung von zwei Fehlern  $A + B$ ; bei  $m$  bzw.  $n$  Messungen sind  $m n$  Summenkombinationen möglich.) Das letzte Fehlergesetz entsteht durch lineare Kombination der Fehlergesetze  $\varphi(x)$  und  $\psi(x)$ . Nun erhebt sich die Frage, wann die folgende Funktionalgleichung gültig sei:

$$\frac{1}{c} \varphi\left(\frac{x}{c}\right) = \frac{1}{a b} \int_{-\infty}^{+\infty} \varphi\left(\frac{u}{a}\right) \varphi\left(\frac{x-u}{b}\right) du.$$

PÓLYA hat gezeigt, daß dies unter sehr allgemeinen Voraussetzungen nur dann zutrifft, wenn  $\varphi(x)$  die Dichte der Normalverteilung bedeutet. Die Normalverteilung liefert demnach ein Beispiel für eine sogenannte stabile Verteilung, deren Typus auch bei Linearkombination erhalten bleibt.

Der zweite Grund für die große Bedeutung der Normalverteilung liegt darin, daß sie bei sogenannten stetigen stochastischen Prozessen, wie sie in der Physik öfters

betrachtet werden, sehr häufig vorkommt, unter sehr allgemeinen Voraussetzungen überhaupt die einzige mögliche Verteilung darstellt. Wenn sich beispielsweise ein Teilchen in aufeinanderfolgenden Schritten vom Nullpunkt aus längs einer Geraden bewegt und  $x_k$  den Zuwachs der Abszisse beim  $k$ -ten Schritt bedeutet, so stellt die totale Abszisse  $x = x_1 + \dots + x_n$  eine stetige stochastische Variable dar, und das Ganze wird als ein stetiger stochastischer Prozeß bezeichnet. Nach den vorausgegangenen Ausführungen ist es verständlich, daß  $x$  sehr häufig einer Normalverteilung gehorcht. Es sind vor allem russische Mathematiker, wie KHINTSCHINE und KOLMOGOROFF, und daneben der Franzose P. LÉVY, welche die stochastischen Prozesse untersuchten und auch enge Beziehungen mit der Theorie gewisser partieller Differentialgleichungen (parabolische und elliptische Typen) entdeckten. Nach der neuesten Entwicklung in den USA. versucht man umgekehrt, partielle Differentialgleichungen als einen stochastischen Prozeß zu interpretieren und mit den modernen Rechenmaschinen numerisch zu behandeln und zu lösen. Tatsächlich öffnen sich damit jetzt neue und vielleicht sehr fruchtbare Möglichkeiten der Anwendung der Wahrscheinlichkeitsrechnung auf die Analysis.

Nach Aussagen namhafter Biologen wäre es sehr wertvoll, wenn unsere schweizerischen Mittelschüler und insbesondere auch die Gymnasiasten eine bescheidene Einführung in die allerelementarsten Fragestellungen und Begriffe der Wahrscheinlichkeitsrechnung und mathematischen Statistik in ihrem Mathematikunterricht erhielten<sup>1)</sup>. Als Mathematiker kann ich mich diesem Standpunkt nur anschließen. Es ergibt sich hier eine neue Möglichkeit, die Mathematik mit andern Wissenschaften zu verknüpfen und die erkenntnistheoretische Tragweite ihrer Methoden von allgemeinen Gesichtspunkten aus zu beleuchten.

Zum Schlusse seien einige Bücher aus der überaus umfangreichen Literatur betreffend Wahrscheinlichkeitsrechnung und mathematische Statistik genannt:

- V. MISES, *Wahrscheinlichkeitsrechnung* (S. Rosenberg, Publishers, New York 1945).  
 H. CRAMÉR, *Mathematical Methods of Statistics* (Princeton 1946).  
 A. KHINTSCHINE, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung* (Springer, Berlin 1933).  
 A. KOLMOGOROFF, *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Springer, Berlin 1933).

WALTER SAXER, Zürich

## Die Steinersche Hypozykloide

*Vorbemerkung:* JAKOB STEINER hat die wichtigsten Eigenschaften dieser Kurve im Jahre 1856 bekanntgegeben, nach seiner Gewohnheit ohne Beweise (Ges. Werke, Bd. 2, S. 641–647). CLEBSCH, CREMONA, BATTAGLINI, PAINVIN, LAGUERRE u. a. haben das Thema aufgegriffen, zum Teil unter allgemeinen Gesichtspunkten. Eine reiche Fülle von Eigenschaften findet man in einer neueren Arbeit von J. LEMAIRE<sup>2)</sup>. Die vorliegende Arbeit schöpft aus diesen Quellen, geht aber zum Teil eigene Wege. Sie will auf den schönen und so leicht erreichbaren Garten hinweisen, indem sie eine Aus-

<sup>1)</sup> Siehe z. B. H. STOHLER, *Algebra*, Leitfaden dritter Teil (Unterrichtswerk des Vereins Schweizerischer Mathematiklehrer) (Orell Füßli Verlag, Zürich 1938).

<sup>2)</sup> J. LEMAIRE, *Hypocycloïdes et Epicycloïdes* (Vuibert, Paris 1929).