

Zeitschrift: Mitteilungen der Naturforschenden Gesellschaft Bern
Herausgeber: Naturforschende Gesellschaft Bern
Band: - (1893)
Heft: 1305-1334

Artikel: Beiträge zur Darstellung des Bernoulli'schen Theorems, der Gammafunktion und des Laplace'schen Integrals
Autor: Eggenberger, J.
DOI: <https://doi.org/10.5169/seals-319064>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. 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. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

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. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

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. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 10.02.2026

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

J. Eggenberger.

Beiträge zur Darstellung des Bernoulli'schen Theorems, der Gammafunktion und des Laplace'schen Integrals.

Eingereicht im August 1893.

Vorbemerkungen.

Die vorliegende Arbeit wurde auf Anregung meines verehrten Lehrers, des Herrn Prof. Dr. J. H. Graf, unternommen.

Sie zerlegt sich in zwei Theile, von denen der erste (die Abschnitte I—VI) historischer, der zweite (die Abschnitte VII und VIII) analytischer Natur ist. Abschnitt I weist einleitend mit einigen Belegen auf den fructificirenden Einfluss der Entwicklung der Wahrscheinlichkeitsrechnung auf diejenige der Analysis hin und präcisirt den Zweck der historischen Untersuchung des ersten Theils. In Abschnitt II wird sodann die philosophische und analytische Begründung des Gesetzes der grossen Zahlen nach Bernoulli's Ars conjectandi gegeben. Die Abschnitte III, IV und V sind den mit Erfolg gekrönten Bemühungen Moivres, dem Bernoulli'schen Theorem einen bestimmten mathematischen Ausdruck zu verleihen, gewidmet, stellen das Summationsverfahren jenes Mathematikers zur Bestimmung eines Näherungswertes für den Binomialcoefficienten dar, beleuchten die Verdienste Moivres und Stirlings um die Darstellung eines Näherungswertes für $\log \Gamma(x)$ und geben die Moivre'sche Darstellung des Laplace'schen Integrals. Abschnitt VI zeigt die Auffindung einer Summationsformel durch MacLaurin und Euler, die in hinreichend allgemeiner Weise gestattet, dem Bernoulli'schen Theorem jenes analytische Gewand zu geben, dessen Schöpfer Laplace ist, da mittelst jener Formel Näherungswerte sowohl für $\log \Gamma(x)$ wie auch für die Summe von Termen einer binomischen Entwicklung von sehr hoher Potenz innerhalb gewisser

Grenzen gefunden werden können. Den Schluss dieses Abschnittes bildet eine Zusammenstellung der gewonnenen historischen Resultate.

Der analytische Theil enthält zunächst (in Abschnitt VII) eine Untersuchung des Verfassers über eine Verallgemeinerung der von J. A. Serret gegebenen, eleganten Entwicklung eines Näherungswertes für $\Gamma(x+1)$ aus der Formel von Wallis, zeigt dann durch eine weitere Untersuchung (in Abschnitt VIII), dass der immer noch gebräuchliche Laplace'sche Ausdruck für das Bernoulli'sche Theorem, nämlich

$$\frac{2}{\pi} \int_0^{\gamma} e^{-t^2} dt + \sqrt{\frac{e^{-\gamma^2}}{2\pi\mu p q}} \text{ gleich ist } \frac{2}{\pi} \int_0^{\delta} e^{-t^2} dt.$$

Diese Vereinfachung des Laplace'schen Ausdrucks dürfte für die Wahrscheinlichkeitsrechnung und die Versicherungstechnik von Werth sein.

In den Anhang wurden neben dem Quellenverzeichniss einige Anmerkungen, die den Text sonst allzu störend unterbrochen hätten, als Noten verwiesen.

I.

1. Seit *Laplace* und *Gauss* ist die Wahrscheinlichkeitsrechnung für die exakte wissenschaftliche Forschung ein unentbehrliches Hilfsmittel geworden und auch bei Fragen der Sozialpolitik und der Kultur im weiteren Sinne ist sie berufen, immer werthvollere Dienste zu leisten. Neben diesem ihrem Anteil an der Entwicklung der beobachtenden Wissenschaften ist aber auch der Gewinn nicht unbedeutend, den diese angewandte mathematische Disciplin der reinen Mathematik gebracht hat. Denn ähnlich wie andere angewandte mathematische Wissenschaften, die Astronomie und die mathematische Physik, auf die Erfindung und Entwicklung der Infinitesimalrechnung und auf die Theorie der partiellen Differentialgleichungen im höchsten Grad anregend gewirkt haben, so ist auch die Wahrscheinlichkeitsrechnung nicht ohne Einfluss auf die Entwicklung der Analysis des Endlichen und Unendlichen gewesen. Ein kurzer Blick in deren Geschichte soll uns davon überzeugen.

Die Wahrscheinlichkeitsrechnung nahm ihren Ursprung im 17. Jahrhundert, in der Zeit der mathematischen Entdeckungen. Einige Würfelspielprobleme, die ihm vom *Marquis de Méré* im Jahre 1654 vorgelegt wurden, veranlassten den geistvollen französischen Philosophen und Mathematiker *Blaise Pascal* (1623—1662) mit der Unter-

stützung seines Zeitgenossen Pierre Fermat (1608—1665) genauer mit dem neuen Calcul sich zu beschäftigen, und die ersten Prinzipien desselben feststellend, wurden Pascal und Fermat die Begründer der Wahrscheinlichkeitsrechnung, von ihnen «géometrie du hasard» auch «aleæ geometria» genannt. Weil aber die Hilfsmittel der Analysis damals für die Lösung der Spielprobleme keine genügenden waren, erweiterte Pascal die Combinationslehre*) und zeigte deren Zusammenhang mit den figurirten Zahlen**).

Der grosse Basler Mathematiker *Jakob Bernoulli* I. (1654—1705) gab dann in seinem epochemachenden Werke über Wahrscheinlichkeit, *Ars conjectandi****) (Muthmassungskunst) eine beinahe vollständige Theorie der Combinatorik, der figurirten Zahlen†) und fand auch die nach ihm benannten Zahlen†), die bekanntlich in der Reihen- und Interpolationstheorie von Wichtigkeit sind.

Pierre Raimond de Montmort (1678—1719) lieferte im Dienste der Wahrscheinlichkeitsrechnung ebenfalls Beiträge zur Analysis der Reihen††), namentlich in Bezug auf die Summation von arithmetischen Reihen höherer Ordnung.

Ein anderer, sehr bedeutender französischer Mathematiker, der nach Aufhebung des Ediktes von Nantes in London ein Asyl gefunden hatte, *Abraham de Moivre*, entdeckte bei seinen Studien über die Wahrscheinlichkeitsrechnung die recurrenten Reihen, deren Theorie er in dem für die Analysis bedeutsamen Buche: *Miscellanea analytica de seriebus et quadraturis* (London 1730) vortrug†††). Moivres weitere sehr werthvolle Beiträge zur Analysis werden im Verlaufe meiner historischen Untersuchung noch deutlicher hervortreten.

Den Forschungen der beiden grossen französischen Analysten, *Joseph Louis Lagrange* (1736—1813) und *Pierre Simon Laplace* (1749—

*) Die Anfänge der Combinatorik waren aus einer Schrift Guldins vom Jahre 1622 bekannt.

**) In einem nachgelassenen Werke Pascals: *Traité du triangle arithmétique*. Paris 1665.

***) Basel 1713. Herausgegeben und mit einem Vorwort versehen von Nikolaus Bernoulli, dem Neffen Jakob Bernoulli's.

†) *Ars conjectandi*, Lib. II.

††) *Montmort, Essai d'analyse sur le jeu de hasard*. Paris 1708.

†††) Lib. II. Cap. II. *De natura serierum recurrentium*.

Lib. IV. Cap. II. *De summis serierum recurrentium*.

Auch Moivres *Doctrine of chances* enthält in der 2. Ausgabe (London 1738) einen Abriss der Theorie von «the summation of the recurring series», p. 193 ff.

bis 1827), auf dem Gebiete der Wahrscheinlichkeit verdankt die höhere Analysis (die sich allerdings inzwischen durch die Arbeiten von Newton, Leibnitz, Moivre, Stirling, Taylor, Mac-Laurin, der Bernoulli, Euler u. a. schon bedeutend entwickelt hatte) ebenfalls neue und wichtige Kapitel.

Schon 1759 veröffentlichte*) der 23jährige Professor an der Artillerieschule in Turin, Lagrange, eine für die Differenzenrechnung epochemachende Abhandlung über «L'intégration d'une équation différentielle à difference finie qui contient la théorie des suites recurrentes», worin die Theorie der recurrenten Reihen verallgemeinert und deren Bedeutung für die Wahrscheinlichkeitsrechnung hervorgehoben wird.

Derjenige, welcher die Bedeutung der Lagrange'schen Arbeit am klarsten erkannte, war der ebenfalls noch junge Professor an der Pariser Militärakademie, Laplace. Schon 1774 schrieb er sein *Mémoire sur les suites recurro-recurrentes et sur leurs usages dans la théorie des hasards.***) In der Vorrede zu einem andern *Mémoire****) sur la probabilité konnte er schreiben: «J'ose me flatter que l'analyse dont je me servis pour cet object pourra mériter l'attention des géomètres». Aus den vielen und langjährigen Arbeiten von Laplace über die Wahrscheinlichkeitsrechnung ging schliesslich sein grosses Werk über diesen Gegenstand, die *Théorie analytique des probabilités*,†) hervor, welches nicht nur für die Wahrscheinlichkeitsrechnung grundlegend, sondern auch für die *Integralrechnung, die Funktionen- und Interpolationstheorie sehr werthvoll ist.*

Die vorstehenden Notizen mögen dargethan haben, wie der Wahrscheinlichkeitsrechnung durch die Auffindung analytischer Hilfsmittel nicht nur die Pfade ihrer eigenen Entwicklung geebnet wurden, sondern wie sie dadurch ihrerseits auch einen wesentlichen fördernden Einfluss auf die Analysis ausgeübt hat.

Als Frucht der Wahrscheinlichkeitsrechnung darf auch das La place'sche Integral, welches in der mathematischen Physik eine grosse Rolle spielt,

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$$

*) In *Miscellanea Taurinensia*, tome I. pag. 33—42.

**) In den «*Mémoires, présentés par divers savants.* t. VI. p. 353—371.

***) *Histoire de l'Académie des sciences pour l'année 1778.* p. 227 ff. Auf den Inhalt dieser Abhandlung soll später noch zurückgekommen werden.

†) Das klassische, Napoleon I. gewidmete Buch, erschien zum ersten Mal anno 1812.

zeichnet werden. Man erhält dieses Integral aus dem Bernoulli'schen Theorem.

Sind p und q die einfachen und konstanten Wahrscheinlichkeiten zweier entgegengesetzter Ereignisse E und E' , so ist die Wahrscheinlichkeit dafür, dass in einer sehr grossen Anzahl von $\mu = m + n$ von Versuchen das Ereigniss E in einer Anzahl von m Malen, wobei m zwischen $\mu p \pm 1$ liegt, eintreffe (vorausgesetzt, dass für ein μp -maliges Eintreffen des Ereignisses E das Maximum von Wahrscheinlichkeit vorhanden), ausgedrückt durch

$$W = \sum_{m=\mu p-1}^{\mu p+1} \frac{\mu!}{m! n!} p^m q^n$$

und zwar kann diese Wahrscheinlichkeit mit wachsendem μ beliebig nahe der Einheit gebracht werden.

Der Summenausdruck kann nun (vermittelst mehrmaliger Näherungen) in folgenden Integralausdruck übergeführt*) werden:

$$W = \frac{2}{\pi} \int_0^{\gamma} e^{-t^2} dt + \frac{e^{-\gamma^2}}{\sqrt{2\pi\mu pq}}$$

Es ist dies ebenfalls die Wahrscheinlichkeit dafür, dass m innerhalb der Grenzen $\mu p \pm 1$ oder hier nun innerhalb $\mu p \pm \gamma \sqrt{2\mu pq}$ liege, wo

$$\gamma = \frac{1}{\sqrt{2\mu pq}}$$

eine Funktion von 1 , μ und p ist.

Den Summenausdruck für W hat Jakob Bernoulli I. schon zu Anfang des vorigen Jahrhunderts gegeben, der Integralausdruck aber in obiger Form wurde erst beinahe ein Jahrhundert später von Laplace aufgestellt.

*Die Festlegung jener Summe durch Jakob Bernoulli, deren Entwickelungsprocess bis zum Integralausdruck und die dabei aufgetretenen analytischen Methoden und Resultate historisch klar zu legen, ist die Aufgabe, die ich im ersten Theil meiner Arbeit zu lösen versucht habe. Dabei waren mir die vortrefflichen Notizen von Laplace**) und*

*) Vrgl. Note 1 im Anhang.

**) Laplace, *Essai philosophique sur les probabilités*, veröffentlicht als Einleitung in der *Théorie analyt. des probabilités* und in einer Separatausgabe.

Todhunter*) über die Geschichte der analytischen Darstellung des Bernoulli'schen Theorems wegleitend.

Im *Essai philosophique sur les probabilités***) sagt Laplace im Abschnitt: les lois de la probabilité qui résultent de la multiplication indéfinie des événemens: «Ce théorème indiqué par le bon sens «était difficile à démontrer par l'analyse. Aussi l'illustre géomètre «Jacques Bernoulli qui s'en est occupé le premier, attachait-il une «grande importance à la démonstration qu'il en a donnée». Weiter im Abschnitt: *Notice historique sur le calcul de probabilité*, wo Laplace von Bernoulli's *Ars conjectandi* spricht, finden wir:***)

«Cet ouvrage est encore remarquable par la justesse et la finesse «des vues, par l'emploi de la formule du binôme dans ce genre de questions, et par la démonstration de ce théorème, savoir, qu'en multipliant «indéfiniment les observations et les expériences; le rapport des événemens de diverses natures, approche de celui de leurs possibilités respectives, dans des limites dont l'intervalle se resserre de plus en plus, en «mesure qu'ils se multiplient et devient moindre qu'aucune quantité assignable. Ce théorème est très utile pour reconnaître par les observations, «les lois et les causes des phénomènes. Bernoulli attachait avec raison, «une grande importance à sa démonstration qu'il dit avoir méditée pendant vingt années.

«Moivre a repris dans son ouvrage le théorème de Jacques Bernoulli «sur la probabilité des résultats déterminés par un grand nombre d'observations. Il ne se contente pas de faire voir comme Bernoulli, que «le rapport des événemens qui doivent arriver, approche sans cesse de «celui de leurs possibilités respectives; il donne de plus une expression «élégante et simple de la probabilité que la différence de ces deux rapports est contenue dans des limites données. Pour cela, il détermine «le rapport du plus grand terme du développement d'une puissance très «élevée du binôme, à la somme de tous ses termes; et le logarithme hyperbolique de l'excès de ce terme, sur les termes qui en sont très voisins. Le plus grand terme étant alors le produit d'un nombre considérable de facteurs; son calcul numérique devient impraticable. Pour «l'obtenir par une approximation convergente, Moivre fait usage d'un théorème de Stirling sur le terme moyen du binôme élevé à une haute puissance, théorème remarquable, surtout en ce qu'il introduit la racine

*) J. Todhunter, *A history of the mathematical theory of probability from the time of Pascal to that of Laplace*. London 1865.

**) Separatausgabe (3. éd. Paris 1816) p. 74; *Théorie analyt. des probabilités, introduction* p. XLVII.

***) L. c. p. 211; p. CXLVI.

«carrée du rapport de la circonférence au rayon, dans une expression «qui semble devoir être étrangère à cette transcendante. Aussi Moivre «fut-il singulièrement frappé de ce résultat que Stirling avait déduit de «l'expression de la circonférence en produits infinis, expression à laquelle «Wallis était parvenu par une singulière analyse qui contient le germe «de la théorie si curieuse et si utile des intégrales définies.»

Den Laplace'schen Bemerkungen zur Geschichte des Bernoulli'schen Theorems lasse ich noch die Uebersicht folgen, die J. Todhunter*) über die nämliche Materie gibt: «With respect to the history of the «result obtained in art. 994 (Laplace'sche Darstellung des Bernoulli- «schen Theorems), we have to, remark that James Bernoulli began «the investigation; then Stirling and De Moivre carried it on by the «aid of the theorem known by Stirling's name; and lastly, the theo- «rem known by Euler's name gave the mode of expressing the finite «summation by means of an integral. But it will be seen that prac- «tically we use only the first term of the series given in Euler's «theorem, in fact no more than amounts to evaluating an integral by «a rough approximate quadrature. Thus the result given by Laplace «was within the power of mathematiciens as soon as Stirling's Theo- «rem had been published.»

Das vortreffliche Werk Todhunders über die Geschichte der Wahrscheinlichkeitsrechnung gibt die Notizen über das Bernoulli'sche Theorem zerstreut bei der Besprechung der Arbeiten von Bernoulli, Moivre und Laplace über die Wahrscheinlichkeitsrechnung. Dagegen konnte in seiner Geschichte der Wahrscheinlichkeitsrechnung auf die Darstellung der analytischen Hilfsmittel desselben gar nicht eingegangen werden. Eine zusammenhängende, eingehende Darlegung dieser Verhältnisse, besonders wenn sie wesentlich neue Resultate zu Tage zu fördern vermag, schien mir daher ebenso interessant wie werthvoll zu sein.

II.

3. In einem Begleitschreiben zu seiner Schrift: *De rationiis in ludo aleae***), schrieb der gelehrte Huygens an seinen Lehrer der Mathematik Franziskus von Schooten u. a. Folgendes:

*) J. Todhunter, *History of the mathematical theory of probability*, art. 995 pag. 553.

**) Diese Arbeit erschien als Anhang zu Schootens *Exercitationes mathematicae*, 1657. Huygens hat darin zum ersten Mal die Prinzipien der Wahrscheinlichkeitslehre systematisch und analytisch formulirt, so dass Jacob Bernoulli diese Huygen'sche Schrift dann in sein erstes Buch der *Ars conjectandi* aufgenommen und commentirt hat.

«Quanquam, si quis penitus ea quae tradimus examinare caeperit,
«non dubito quin continuò reperturus sit, rem non, ut videtur, ludicram
«agi, sed pulchrae subtilissimaeque contemplationis fundamenta explicari.
«Et problemata quidem, quae in hoc genere proponuntur, nihilo minus
«profundae indaginis visum iri confido, quam quæ Diophanti libris conti-
«nentur, voluptatis autem aliquanto plus habitura, cum non, sicut illa,
«in nuda numerorum consideratione terminentur.»

Bekundet damit Huygens eine hohe Meinung von der Wichtigkeit des neuen Calcüls und verheisst er demselben eine grosse Zukunft, so gelang es ihm aber doch noch nicht, sich über das Niveau der üblichen Anwendung der Wahrscheinlichkeitstheorie, die sich bis zu jener Zeit auf das Gebiet der Spielprobleme beschränkt hatte, zu erheben.

Wenige Jahre später machte zwar der berühmte Grosspensionär von Holland, Jean de Witt, der treffliche Kenner und Förderer der Cartesianischen Geometrie, die ersten nützlichen Anwendungen auf die Rentenrechnung*); aber es blieb dem genialen Kopfe Jakob Bernoulli's I. vorbehalten, der neuen mathematischen Disciplin ihr weites Arbeitsfeld zu eröffnen.

In einer Zeit grosser wissenschaftlicher Entdeckungen hatte sich Bernoulli's schöpferische Kraft entfaltet. Längst schon hatten Bacon von Verulam, Giordano Bruno u. a. m. der wissenschaftlichen Forschung den Weg der Beobachtung gewiesen und eine Reihe von grossen Forschern hatte bereits die neue Methode der Induction durch glänzende Erfolge gerechtfertigt. Kopernikus hatte die richtige Vorstellung von unserem Planetensystem gegeben, Kepler seine Gesetze der Planetenbewegung berechnet, Galilei die Fallgesetze erkannt und Newton der letzteren Gültigkeit im Universum als Gravitationsgesetz nachgewiesen. Vieles, was früher als zufällig erscheinen mochte, war durch Causalgesetze erklärt und die Domäne des Zufalls und des Aberglaubens hatte schon bedeutend an Terrain verloren. Und dennoch waren es kühne Fragen, die Bernoulli's weiter Blick in den Thatsachen zu lesen vermochte. Gibt es in den gesammten Erscheinungen überhaupt einen Zufall? Erscheint uns vielleicht das *anscheinend Zufälligste* nur *deshalb zufällig*, weil wir *seine Ursachen nicht zu ergründen vermögen*? Ist es möglich, durch fortgesetzte Beobachtungen auch das Zufälligste

*) Jean de Witt, *De vardye van de lifrenten na proportie van de losrenten, ou la valeur des rentes viagères en raison des rentes libres et remboursables.* La Haye 1671.

als von Gesetzen abhängig zu erkennen? Ist es überhaupt möglich, durch Beobachtungen ein genügend sicheres Resultat zu erhalten? Und in welcher Beziehung steht die Zahl der Beobachtungen zur Genauigkeit des Resultates?

4. Jakob Bernoulli I. hat seine diesbezüglichen Gedanken in dem hochinteressanten vierten Buche seiner *Ars conjectandi**, betitelt: *Ad Usum et applicationem praecedentis Doctrinae in Civilibus, Moralibus et Oeconomicis*, niedergelegt. Das nach ihm benannte Theorem**) findet sich dort im 4. und 5. Kapitel. Die Hauptgedanken sollen ihrer grundlegenden Bedeutung wegen hier ihre Stelle finden. Cap. IV. betitelt: *De dupli Modo investigandi numeros casum. Qui sentiendum de illo, qui instituitur per experimenta. Problema singulare eam in rem propositum*, hat zusammengefasst folgenden Inhalt:

Es wurde im letzten Cap. (III) gezeigt, wie die Beweiskraft von Argumenten für gewisse Dinge nach der Zahl von günstigen und ungünstigen Fällen durch Rechnung zu schätzen ist. Hier aber liegt die Schwierigkeit; denn nur für die wenigsten Erscheinungen ist die Zahl der günstigen oder ungünstigen Fälle und das Gewicht jedes Einzelnen bekannt. Beim Würfelspiel ist es allerdings nicht schwer, die Zahl der günstigen Fälle für das Eintreffen eines bestimmten Ereignisses zu berechnen und ebenso leicht ist es, die Fälle für das Ziehen eines weissen oder schwarzen Steinchens aus einer Urne, wenn das Verhältniss der verschiedenartigen Steinchen gegeben ist, zu bestimmen. Wer könnte aber jemals die Anzahl von Krankheiten, die den menschlichen Körper an allen Theilen und zu jedem Alter befallen und den Tod herbeiführen können, bestimmen und herausfinden, um wie viel leichter diese oder jene Krankheit den Tod herbeiführen können, so dass dann eine Vermuthung über das Leben eines Menschen oder dasjenige zukünftiger Generationen ausgesprochen werden könnte? Oder wer könnte die zahllosen Fälle von Veränderungen ergründen, denen die Luft tagtäglich ausgesetzt ist, um heute schon Vermuthungen über deren Zustand nach einem Monat oder nach einem Jahr aufzustellen? Oder wer kennt die Natur des menschlichen Geistes und den wunderbaren Bau unseres Körpers so genau, dass er bei einem Spiele, das grösstentheils von der Schnelligkeit und dem Verstande des Spielers abhängt, die Fälle vorauszusagen sich unterstünde, in welchen dieser oder jener Spieler gewinnt oder verliert?

*) Von der Liagre in seinem *Calcul des probabilités* sagt: «Cette ouvrage contient en germe toute la philosophie de la probabilité».

**) «The memorable theorem in the fourth part, which justly bears its authors name, will ensure him a permanent place in the history of the Theory of Probability.» J. Todhunter, *History of the Theory of Probability* p. 77.

Wegen der Beschränktheit unseres Geistes wäre es also ein eitles Bemühen *die verschiedenen Fälle a priori auffinden zu wollen; doch steht uns hier der Weg der Beobachtung offen: wir können die Wahrscheinlichkeit auch a posteriori, durch Beobachtung finden. Voraussetzung ist dabei, dass für bestimmte Ereignisse eine gewisse Konstanz der Ursachen angenommen werde.* Denn, wenn z. B. einmal 300 Menschen untersucht worden sind vom Alter und der Konstitution des Titius und man gefunden hätte, dass 200 davon vor Verfluss von 10 Jahren gestorben sind, so kann man den Schluss ziehen, dass es 2 Mal mehr Fälle gibt dafür, dass auch Titius innerhalb von 10 Jahren sterben, als dass er diesen Zeitraum überleben werde. Ebenso wenn einer mehrere Jahre das Wetter beobachtet, wenn er oft bei 2 Spielenden gestanden und deren Spiel verfolgt hat, so kann er mit ziemlicher Sicherheit die Wahrscheinlichkeit bestimmen dafür, dass ein diesbezügliches Ereigniss unter denselben Umständen eintritt oder nicht eintritt.

Und diese empirische Art der Bestimmung der Zahl von Fällen durch Beobachtungen ist weder neu noch ungewohnt und wird in der Praxis von jedermann angewendet. Auch ist jedem klar, *dass um einen richtigen Schluss ziehen zu können, nur wenige Beobachtungen nicht genügen, sondern dass eine grosse Anzahl derselben nöthig sind.* Obgleich diess nun aber aus der Natur der Sache von jedermann eingesehen wird, so liegt doch der auf wissenschaftlichen Prinzipien gegründete Beweis durchaus nicht auf der Oberfläche. *Es muss vielmehr untersucht werden, was vielleicht noch niemand eingefallen ist, ob durch Vermehrung der Beobachtungen auch die Wahrscheinlichkeit vermehrt werde dafür, dass die Zahl der günstigen zu den ungünstigen Beobachtungen ein wahres Verhältniss erreiche und dass diese Wahrscheinlichkeit zuletzt jeden beliebigen Grad von Gewissheit erreichen könne, oder ob das Problem vielmehr, um so zu sagen, seine Asymptoten hat, d. h. ob ein bestimmter Grad der Gewissheit gegeben sei, der auch bei beliebiger Vermehrung der Beobachtungen niemals überschritten werden könnte, z. B. $\frac{1}{2}$ oder $\frac{2}{3}$ oder $\frac{3}{4}$ der Gewissheit.* Seien z. B. in einer Urne ohne dein Wissen 3000 weisse und 2000 schwarze Steinchen verborgen und du nehmest, um das Verhältniss derselben zu bestimmen, ein Steinchen nach dem andern heraus (so jedoch, dass du das gezogene, bevor du ein neues ziehst, wieder hineinlegst), und du beobachtest nun, wie oft ein weisses, wie oft ein schwarzes herauskommt. Die Frage ist nun, wie oft du dies thun kannst, damit es 10-, 100-, 1000-fach wahrscheinlicher (d. h. am Ende intellectuell gewiss) werde, dass die Zahl der Male, in denen du ein weisses, zu denen, in welchen du ein schwarzes bekommst, das Verhältniss $1\frac{1}{2}$ bilde, als dass dieses Verhältniss ein anderes davon verschiedenes sei. Ist dies nicht der Fall, so ist unser Unternehmen, die Zahl der Fälle durch Versuche zu bestimmen, werthlos. Wenn es aber der Fall ist (was wir im folgen-

den Cap. [V] zeigen werden), so können wir die Zahl der Fälle a posteriori erforschen, wie wenn sie uns a priori bekannt wären und das ist im praktischen Leben, wo das der Vernunft Gewisse als absolut gewiss angesehen wird, genügend, um unsere Vermuthungen in einem beliebigen Zufallsgebiet nicht weniger wissenschaftlich zu leiten als bei den Würfelspielen. Denn stellen wir uns vor, dass die Luft oder der menschliche Körper den Herd vieler Veränderungen und Krankheiten in sich schliessen, gerade so wie die Urne die Steinchen, so werden wir ebenfalls auf diesem Gebiet bestimmen können, wie viel leichter dieses oder jenes Ereigniss eintreten kann als ein anderes.

Es ist noch zu bemerken, dass ich das Verhältniss der durch die Beobachtung zu bestimmenden Fälle nicht ganz genau angeben, sondern in gewisse Grenzen einschliessen will. Im oben gegebenen Beispiel würden wir vielleicht das Verhältniss $1\frac{1}{2}$ einschliessen zwischen $\frac{301}{200}$ und $\frac{299}{200}$ oder zwischen $\frac{3001}{2000}$ und $\frac{2999}{2000}$. Es zeigt sich dann, dass es durch fortgesetzte Beobachtungen immer wahrscheinlicher wird, dass das durch Beobachtung gefundene Verhältniss der Fälle innerhalb, als dass es ausserhalb dieser Grenzen liegt.

Jakob Bernoulli schliesst den Kommentar zu seinem Theorem wörtlich so: «Hoc igitur est illud Problema, quod evulgandum hoc «loco proposui, postquam jam per vicennium pressi, et cuius per novi- «tas, tum summa utilitas cum pari conjuncta difficultate omnibus reli- «quis hujus doctrinae capitibus pondus et pretium superaddere potest.»

Schliesslich wendet sich Jakob Bernoulli noch polemisirend an gewisse Gelehrte*), welche gegen seine Theorie Einwände zu machen versucht hatten.

1) Werfen sie vor, das Verhältniss zwischen den Steinchen sei anders beschaffen als dasjenige zwischen den Krankheiten oder den Luftveränderungen; die Zahl jener sei bestimmt, die Zahl dieser dagegen unsicher und unbestimmt. Antwort: Beides ist nach unserer Erkenntniss gleich unsicher und gleich unbestimmt; aber das was an sich oder von Natur aus so ist, dass es von uns nicht allseitig erkannt werden kann, dasselbe ist ebenfalls von Gott erschaffen, und was Gott erschaffen, das bestimmte er auch, ehe er es schuf.

2) Bemerken sie: die Zahl der Steinchen sei endlich, die der Krankheiten aber nicht. Antwort: Sie ist eher erstaunlich gross als unendlich; aber zugegeben, sie sei unendlich, so ist bekannt, dass auch zwischen zwei unendlichen Grössen ein bestimmtes Verhältniss bestehen kann und dass dasselbe auch durch endliche Grössen genau oder wenigstens an-

*) Es ist damit wohl Leibnitz gemeint, der über diesen Gegenstand in Briefen an Bernoulli polemisirte.

nähernd bestimmt werden kann. Ich erinnere z. B. an die Ludolf'sche Zahl. *Es hindert daher nichts, dass ein Verhältniss zwischen unendlichen Grössen doch durch eine endliche Zahl annäherungsweise ausgedrückt und durch eine endliche Zahl von Beobachtungen bestimmt werden kann.*

3) Wenden sie ein: *die Zahl der Krankheiten sei nicht constant, sondern täglich entstünden neue.* Antwort: Dass sich im Laufe der Zeit die Krankheiten vermehren, kann man nicht läugnen, *und sicherlich wird der, welcher aus heutigen Beobachtungen auf antediluviale Zeiten schliessen wollte, sehr irren.* Aber hieraus folgt nur, dass bisweilen neue Beobachtungen zu machen sind, wie sie bei den Steinchen zu machen wären, wenn die Vermuthung nahe läge, dass sich ihre Zahl verändert hätte.

5. Im V. Kapitel: «Solutio Problematis praecedentis», gibt Jakob Bernoulli I. die analytische Darstellung seines Theorems wie folgt*):

Lemma I.

Sei gegeben die Reihe

0, 1, 2, $r - 1, r, r + 1, \dots, r + s - 1, r + s$ und es werde dieselbe fortgesetzt bis ihr letztes Glied $nr + ns$ heisst, so entsteht die neue Reihe

0, 1, 2, $nr - n, \dots, nr, \dots, nr + n, \dots, nr + ns$, in welcher die Zahl der Glieder zwischen $nr + n$ und $nr + ns$ die Gliederzahl zwischen nr und $nr + n$ nicht mehr (wie gross auch n werde) als $s - 1$ mal übertrifft und die Zahl der Glieder links von $nr - n$ die Zahl der Glieder zwischen $nr - n$ und nr nicht mehr als $r - 1$ mal.

Lemma II. Wenn das Binom $(r + s)$ in irgend eine Potenz erhoben wird, so hat die Entwicklung immer ein Glied mehr als der Exponent Einheiten.

Lemma III. In der Entwicklung von $(r + s)^n$ ist ein Term M dann der grösste, wenn die Zahl der vorausgehenden Glieder zur Zahl der nachfolgenden, mit r und s , in indirekter, oder wenn die Dimensionen von r und s in M mit r und s in direkter Proportion stehen.

Dieser Term M hat zum näheren einen kleineren Verhältnisswerth als — bei gleichem Intervall — der nähere zum entfernteren.

Demonstr. 1. Setzt man $nt = nr + ns$, so wird

$$(r + s)^{nt} = r^{nt} + \binom{nt}{1} r^{nt-1} \cdot s + \dots + \binom{nt}{nt-1} r \cdot s^{nt-1} + s^{nt}$$

und der grösste Term

*) In gedrängter Uebersicht.

$$M = \frac{nt(nt-1)(nt-2)\dots(nt-ns+1)}{1 \cdot 2 \cdot 3 \dots ns} r^{nr} \cdot s^{ns} \text{ oder}$$

$$M = \frac{nt(nt-1)(nt-2)\dots(nt-nr+1)}{1 \cdot 2 \cdot 3 \dots nr} r^{nr} \cdot s^{ns}$$

Bezeichnet man ferner mit $R_1, R_2, R_3 \dots$ die rechts von M aus aufeinanderfolgenden Terme, und mit L_1, L_2, L_3, \dots die entsprechenden links, so ist

$$R_1 = \frac{nt(nt-1)(nt-2)\dots(ns+2)}{1 \cdot 2 \cdot 3 \dots (nr-1)} r^{nr-1} \cdot s^{ns+1}$$

$$L_1 = \frac{nt(nt-1)(nt-2)\dots(nr+2)}{1 \cdot 2 \cdot 3 \dots (ns-1)} r^{nr+1} \cdot s^{ns-1}$$

$$R_2 = \frac{nt(nt-1)(nt-2)\dots(ns+3)}{1 \cdot 2 \cdot 3 \dots (nr-2)} r^{nr-2} \cdot s^{ns+2}$$

$$L_2 = \frac{nt(nt-1)(nt-2)\dots(nr+3)}{1 \cdot 2 \cdot 3 \dots (ns-2)} r^{nr+2} \cdot s^{ns+2},$$

woraus sich durch Division ergibt:

$$\left| \begin{array}{l} \frac{M}{L_1} = \frac{(nr+1)s}{n \cdot r \cdot s} \\ \frac{M}{R_1} = \frac{(ns+1)r}{n \cdot r \cdot s} \end{array} \right| \left| \begin{array}{l} \frac{L_1}{L_2} = \frac{(nr+2)s}{(ns-1)r} \\ \frac{R_1}{R_2} = \frac{(ns+2)r}{(nr-1)s} \end{array} \right.$$

Es leuchtet aber ein, dass

$$\begin{array}{l|l} nrs + s > nrs & nrs + 2s > nrs - r \\ nrs + r > nrs & nrs + 2r > nrs - s; \end{array}$$

also ist auch

$$M > R_1, M > L_1, L_1 > L_2, R_1 > R_2.$$

Demonstr. 2. Weil

$$\frac{nr+1}{ns} < \frac{nr+2}{ns-1}, \frac{ns+1}{nr} < \frac{ns+2}{nr-1},$$

so folgt auch

$$\frac{(nr+1)s}{nrs} < \frac{(nr+2)s}{(ns-1)r}, \quad \frac{(ns+1)r}{nrs} < \frac{(ns+2)r}{(nr-1)s}$$

oder

$$\frac{M}{L_1} < \frac{L_1}{L_2}, \quad \frac{M}{R_1} < \frac{R_1}{R_2} \cdot \text{q. e. d.}$$

Lemma IV. In der Potenz eines Binoms, dessen Exponent nt sei, kann n so gross genommen werden, dass der grösste Term M in Bezug auf 2 Terme L und R , welche um das Intervall von n Termen

nach links und rechts von M abstehen, einen grösseren Verhältnisswerth hat, als irgend ein gegebenes Verhältniss.

Demonstr. Es wurde gefunden

$$M = \frac{nt (nt - 1) (nt - 2) \dots (nr + 1)}{1 \cdot 2 \cdot 3 \dots ns} r^{nr} s^{ns}$$

$$= \frac{nt (nt - 1) (nt - 2) \dots (ns + 1)}{1 \cdot 2 \cdot 3 \dots nr} r^{nr} s^{ns},$$

und weil

$$L_n = \frac{nt (nt - 1) (nt - 2) \dots (nr + n + 1)}{1 \cdot 2 \cdot 3 \dots ns - n} r^{nr+n} s^{ns-n},$$

$$R_n = \frac{nt (nt - 1) (nt - 2) \dots (ns + n + 1)}{1 \cdot 2 \cdot 3 \dots nr - n} r^{nr-n} s^{ns+n},$$

so wird

$$\frac{M}{L_n} = \frac{(nr + n) (nr + n - 1) \dots (nr + 1)}{(ns - n + 1) (ns - n + 2) \dots ns} \cdot \frac{s^n}{r^n},$$

$$\frac{M}{R_n} = \frac{(ns + n) (ns + n - 1) \dots (ns + 1)}{(nr - n + 1) (nr - n + 2) \dots ns} \cdot \frac{r^n}{s^n}.$$

Hieraus erhält man, wenn man die Potenzen r^n und s^n auf die einzelnen Faktoren vertheilt,

$$\frac{M}{L_n} = \frac{(nrs + ns) (nrs + ns - s) \dots (nrs + s)}{(nrs - nr + r) (nrs - nr + 2r) \dots nrs},$$

$$\frac{M}{R_n} = \frac{(nrs + nr) (nrs + nr - 2) \dots (nrs + r)}{(nrs - ns + s) (nrs - ns + 2s) \dots nrs}.$$

Dividirt*) man durch n , so folgt für $\lim n = \infty$

$$\frac{M}{L_n} = \frac{(rs + s) (rs + s - r) \dots (rs + s) rs}{(rs - r) (rs - r - r) \dots (rs - s) rs}$$

$$\frac{M}{R_n} = \frac{(rs + r) (rs + r - 2) \dots (rs + r) rs}{(rs - s) (rs - s - r) \dots (rs - s) rs}.$$

Der Werth dieser Quotienten ist aber wegen der unendlichen Anzahl von Factoren, von denen jeder grösser als 1 ist, selber unendlich gross. Wenn aber sowohl $\frac{M}{L_n}$ wie auch $\frac{M}{R_n}$ unendlich gross

*) Bernoulli gebraucht hier bei der analytischen Erläuterung für das Zeichen \pm das wohl bei keinem andern Mathematiker angewendete Zeichen \wp . Sein Gleichheitszeichen ist übrigens immer ∞ .

werden kann, so ist gezeigt, dass in der That der Werth des Verhältnisses vom grössten Term einer binomischen Entwicklung zu einem andern Term grösser ist, als bei irgend einem gegebenen Verhältniss.

Lemma V. *Es kann die Zahl n so gross genommen werden, dass die Summe aller Glieder in der binomischen Entwicklung, genommen vom grössten M nach beiden Seiten bis und mit L_n und R_n , zur Summe aller übrigen Glieder ein Verhältniss von grösserem Werth bildet als irgend ein gegebenes.*

Demonstr. Man bezeichne die Terme links von M wie früher mit $L_1, L_2, L_3 \dots \dots \dots$ links von L_n mit $L_{n+1}, L_{n+2}, L_{n+3}, \dots \dots \dots$, dann ist noch Lem. III.:

$$\frac{M}{L_1} < \frac{L_n}{L_{n+1}}, \frac{L_1}{L_2} < \frac{L_{n+1}}{L_{n+2}}, \frac{L_2}{L_3} < \frac{L_{n+2}}{L_{n+3}}, \dots \dots \dots$$

ebenso

$$\frac{M}{L_n} < \frac{L_1}{L_{n+1}} < \frac{L_2}{L_{n+2}} < \frac{L_3}{L_{n+3}} < \dots \dots \dots$$

Für $\lim n = \infty$ wird nach Lem. IV $\frac{M}{L_n} = \infty$, umso mehr $\frac{L_1}{L_{n+2}} = \infty$ und $\frac{L_2}{L_{n+2}} = \infty, \dots \dots \dots$ Daher schliesslich:

$$\frac{L_1 + L_2 + L_3 + \dots \dots \dots}{L_{n+1} + L_{n+2} + L_{n+3} + \dots \dots \dots} = \infty,$$

d. h. die Summe aller Terme zwischen M und L_n genommen, ist unendlich mal grösser als die Summe von ebenso viel Termen ausserhalb von L_n . Nach Lemma I ist aber die Anzahl der Glieder ausserhalb von L_n $s-1$, also eine endliche Zahl mal grösser als die Anzahl der Glieder zwischen L_n und M ; daher ist die Summe der Glieder zwischen L und M (auch mit Ausschluss von M) unendlich mal grösser als die Summe der Glieder ausserhalb L_n .

Das Nämliche kann gezeigt werden vom Verhältniss der Summe der Glieder zwischen M und R_n zu der Summe derjenigen ausserhalb R_n . Schliesslich wird somit die Summe aller Glieder zwischen L_n und R_n (inclus. L_n, R_n und M) das Unendlichvielfache aller übrigen Glieder.

Scol. Es soll noch gezeigt werden, dass auch dann, wenn n endlich bleibt, die Summe der Terme zwischen L_n und R_n zur Summe der übrigen Terme ein Verhältniss ausmacht, das jedes gegebene Verhältniss C an Werth übertrifft.

Es werde das Verhältniss $\frac{r+1}{r}$, welches kleiner ist als $\frac{rs+s}{rs-r}$, in die m^{te} Potenz erhoben, so dass

$$\left(\frac{r+1}{r}\right)^m \geq c(s-1).$$

Um m zu bestimmen, hat man

$$m \log(r+1) - m \log r \geq \log c(s-1), \text{ also}$$

$$m \geq \frac{\log c(s-1)}{\log(r+1) - \log r}.$$

In Lemma IV wurde das Verhältniss $\frac{M}{L_n}$ aus dem Produkt

$$\frac{nrs+ns}{nrs-ns+r} \cdot \frac{nrs+ns-s}{nrs-nr+2r} \cdots \frac{nrs+s}{nrs} \text{ gefunden.}$$

Wird nun n richtig gewählt, so muss einmal einer dieser Brüche gleich $\frac{r+1}{r}$ sein. Bezeichnen wir die Ordnung dieses Bruches in der Faktorenreihe mit m , so ist

$$\begin{aligned} \frac{r+1}{r} &= \frac{nrs+ns-ms+s}{nrs-nr+mr} \text{ und} \\ n &= m + \frac{ms-s}{r+1}, \\ nt &= mt + \frac{mst-st}{r+1}. \end{aligned}$$

nt ist der Exponent, welcher dem Binom gegeben werden muss, damit der grösste Term M der Entwicklung die Grenze L_n um mehr als $c(s-1)$ übertrifft. Der Beweis ergibt sich so: Der Bruch von der Ordnung m wird durch obige Annahme von n gleich $\frac{r+1}{r}$. Nun ist

aber nach Voraussetzung $\left(\frac{r+1}{r}\right)^m \geq c(s-1)$. Weil nun aber alle Brüche, die in obigem Product dem Factor von der Ordnung m vorausgehen, grösser sind als $\frac{r+1}{r}$, die nachfolgenden aber nach der Einheit convergiren, so muss das Product aller grösser sein als $\left(\frac{r+1}{r}\right)^m$ und also um so mehr grösser als $c(s-1)$. Da nun aber jenes

Product gleich dem Verhältniss von $\frac{M}{L}$ ist, so folgt

$$M > c(s-1)L.$$

Ferner ist

$$\frac{M}{L_n} < \frac{L_1}{L_{n+1}} < \frac{L_2}{L_{n+2}} < \dots < \frac{L_n}{L_{2n}},$$

also

$$\frac{L_1 + L_2 + L_3 + \dots}{L_{n+1} + L_{n+2} + L_{n+3} + \dots} > c(s-1).$$

Weil aber die Gliederzahl ausserhalb L_n ($s-1$) mal grösser ist als diejenige zwischen L und M , so folgt, dass das Verhältniss der Summe der Glieder innerhalb von M und L_n zur Summe aller Glieder ausserhalb von L_n grösser als c ist.

Für die Terme rechts von M erhält man dasselbe Resultat. Ausgehend vom Verhältniss $\frac{s+1}{s} < \frac{rs+r}{rs}$ erhalte ich analog durch dieselbe Betrachtung

$$m \geq \frac{\log c (r-1)}{\log(s+1) - \log s} \text{ und}$$

$$nt = mt + \frac{mrt - rt}{s+1}.$$

Die gestellte Aufgabe ist somit gelöst; es kann eine bestimmte Potenz berechnet werden, welche die verlangte Eigenschaft besitzt.

6. *Propos. Princip.* Es folgt endlich der Satz selbst, zu dessen analytischer Darstellung die vorausgegangenen Lemmata gegeben werden mussten.

Es seien einem Ereigniss r Fälle günstig, s Fälle ungünstig, so dass das Verhältniss der günstigen zu den ungünstigen Fällen genau oder annäherungsweise gleich $\frac{r}{s}$ ist; dann ist das Verhältniss der günstigen zu allen möglichen Fällen, — wenn $r+s=t$ — gegeben durch $\frac{r}{t}$, gelegen zwischen den Grenzen $\frac{r+1}{t}$ und $\frac{r-1}{t}$.

Es ist nun zu zeigen, dass so viele Beobachtungen gemacht werden können, dass es irgend eine beliebige Grösse (etwa c) mal wahrscheinlicher wird, es sei das Verhältniss der günstigen zu allen Beobachtungen innerhalb der Grenzen $\frac{r+1}{t}$ und $\frac{r-1}{t}$, als ausserhalb derselben gelegen.

Demonstr. Angenommen nt sei die Zahl der gemachten Beobachtungen. Dann ist, da nach Voraussetzung jeder Beobachtung r Fälle günstig, s Fälle ungünstig sind, die Wahrscheinlichkeit dafür, dass alle Beobachtungen, oder alle mit Ausnahme von einer, von zweien von dreien etc. ein günstiges Resultat liefern, gegeben resp. durch (Part. I, Prop. XIII.)

$$\frac{r^{nt}}{t^{nt}}, \binom{1}{nt} \cdot \frac{r^{nt-1} s}{t^{nt}}, \binom{nt}{2} \frac{r^{nt-2} s^2}{t^{nt}},$$

$$\binom{nt}{3} \frac{r^{nt-3} s^3}{t^{nt}}, \dots \dots \dots$$

Es sind dies die Glieder der binomischen Entwicklung von $\left(\frac{r+s}{t}\right)^{nt}$. Hieraus ist leicht zu schliessen, dass der Wahrscheinlichkeitsgrad*) dafür, dass das Ereigniss bei nt Versuchen nr mal eintreffe, ns mal nicht, gleich ist dem grössten Terme in der Entwicklung von $(r+s)^{nt}$; ebenso wird die Zahl der günstigen Fälle für das $nr + n$ resp. $nr - n$ malige Eintreffen des Ereignisses bei nt Versuchen gegeben durch die Glieder L_n resp. R_n jener binomischen Entwicklung. *Folglich wird der Wahrscheinlichkeitsgrad dafür, dass das Ereigniss bei einer Zahl von nt Versuchen höchstens $nr + n$ und wenigstens $nr - n$ mal eintreffe, ausgedrückt sein durch die Summation aller Terme innerhalb L_n und R_n .* Der Wahrscheinlichkeitsgrad aber dafür, dass das Ereigniss mehr oder weniger als $nr \pm n$ mal eintreffe, wird ausgedrückt sein durch die Summe aller übrigen Terme, die ausserhalb von L_n und R_n liegen. Da nun aber die Potenz des Binoms so gross genommen werden kann, dass die Summe der Glieder zwischen den Grenzen L_n und R_n mehr als c mal grösser ist als die Summe der übrigen Glieder, so folgt auch, *dass so viele Beobachtungen gemacht werden können, dass der Wahrscheinlichkeitsgrad dafür, dass das Verhältniss der Zahl der günstigen Beobachtungsresultate zur Zahl aller innerhalb der Grenzen $\frac{nr+n}{nt}$ und $\frac{nr-n}{nt}$ oder $\frac{r+1}{t}$ und $\frac{r-1}{t}$ liege, mehr als c mal*

*) Unter dem Wahrscheinlichkeitsgrad eines Ereignisses versteht Bernoulli immer die Zahl der dem betreffenden Ereigniss günstigen Fälle.

den Wahrscheinlichkeitsgrad dafür übertrifft, dass jenes Verhältniss ausserhalb der angegebenen Grenzen liege, mit andern Worten, dass es mehr als c mal wahrscheinlicher wird, es liege die Zahl der günstigen Beobachtungsresultate innerhalb der Grenzen $nr \pm n$ als ausserhalb.

Bei der speciellen Betrachtung erklärt es sich von selbst, dass je grösser r , s und t genommen werden, desto enger die Grenzen $\frac{r+1}{t}$ und $\frac{r-1}{t}$ zusammenrücken, so dass das Verhältniss $\frac{r}{t}$ um so bestimmter gegeben werden kann. Wenn daher das Verhältniss der günstigen zu den ungünstigen Fällen etwa gleich $\frac{3}{2}$ ist, so setze man für r und s nicht 3 und 2, sondern 30 und 20, also $t = 50$, so dass die Grenzen $\frac{31}{50}$ und $\frac{29}{50}$ werden und wenn $c = 1000$ gesetzt wird, so ergibt sich (nach Scol.) als Versuchszahl

links von M

$$m > \frac{\log(c(s-1))}{\log(r+1) - \log r} = \frac{4 \cdot 2787536}{142405} < 301,$$
$$nt = mt + \frac{mst - st}{r+1} < 24728;$$

rechts von M

$$m > \frac{\log c(r-1)}{\log(s+1) - \log s} = \frac{4 \cdot 4623980}{211898} < 211,$$
$$nt = mt + \frac{mrt - rt}{s+1} < 25500.$$

Aus diesem Exempel geht hervor, dass es bei 25500 viel mehr als 1000 Mal wahrscheinlicher ist, dass das Verhältniss der günstigen Beobachtungen zu allen innerhalb die Grenzen $\frac{31}{50}$ und $\frac{29}{50}$ fallen werde als ausserhalb. Und ebenso, wenn man $c = 10,000$ setzt, dass dies mehr als 10,000 mal wahrscheinlicher wird bei 31,258 Experimenten und mehr als 100,000 mal bei 36,966 Experimenten; auf diese Weise kann man in infinitum fortfahren, indem man fortwährend zu 25,500 ein Vielfaches von 5708 addirt. Dann sagt Bernoulli weiter: «Und tandem hoc singulare sequi videtur, quod si eventum omnium observationes per totam aeternitatem continuaretur, — probabilitate ultimo in perfectam certitudinem abeunte — omnia in mundo certis rationibus et constanti vicissitudinis lege contingere deprehenderentur; adeo ut etiam in maxime casualibus atque fortuitis quandam quasi necessitatem,

«et, ut sic dicam, fatalitatem agnoscere teneamur; quam nescio annon
«ipse jam Plato intendere voluerit, suo de universali rerum apocatastasi
«dogmate, secundum quod omnia innumerabilium seculorum decursum
«in pristinum reversura statum praedixit.»

Mit dieser weitausschauenden philosophischen Betrachtung schliesst Jakob Bernoulli I. seine *Ars conjectandi*, das Produkt zwanzigjähriger Geistesarbeit, sein bleibendes Denkmal in der Geschichte der Wahrscheinlichkeitsrechnung.

7. Die neuen genialen Ideen Bernoulli's konnten nicht verfehlten, die Polemik der einen, die Bewunderung der andern Gelehrten hervorzurufen, und es ist dafür nicht uninteressant, was Montmort schrieb:*)
«On ne nous a point appris quels sont les Jeux dont cet Auteur —
«Bernoulli — détermenoit les partis, ni quels sujets de politique et
«de morale il avoit entrepris d'éclaircir, mais quelque surprenant que
«soit ce projet, il y a lieu de croire que ce savant Auteur l'auroit
«parfaitement exécuté. M. Bernoulli étoit trop supérieur aux autres
«pour vouloir en imposer, il étoit de ce petit nombre d'hommes qui
«sont propres à inventer et je me persuade qu'il auroit tenu tout ce
«que promettoit le titre de son livre.»

Bernoulli hat nicht versucht, einen bestimmten mathematischen Ausdruck für die Wahrscheinlichkeit, dass die Zahl der günstigen Beobachtungen innerhalb gewisser Grenzen liege, aufzustellen. Sein sehr allgemeiner aber klarer Beweis bezweckte nur, auf exaktem analytischem Wege festzulegen, dass in der That mit der Vermehrung der Beobachtungen auch die Wahrscheinlichkeit immer grösser und schliesslich zur Gewissheit wird, dass die Zahl der günstigen zu den ungünstigen Beobachtungen dem wahren Verhältniss der für das Ereigniss günstigen zu den ungünstigen Fällen gleich kommt (Gesetz der grossen Zahlen). Schon daraus geht hervor, was Bernoulli übrigens auch ausspricht, wenn er sagt:**) «Nisi enim hoc fiat, fateor actum fore de nostro conatu explorandi numeros casuum per experimenta», *dass er das bewiesene Theorem nur als Hülfsatz für die Erforschung der Wahrscheinlichkeit a posteriori betrachtet.* Und dies möchte ich

*) Montmort, *Essai d'analyse sur les Jeux de hasard*, 1. éd. (Paris 1708) Vorrede p. 6. Montmort kannte die Ideen Bernoulli's, dessen Werk noch nicht erschienen war, aus Fontenelle's *Eloge de Mr. Bernoulli*, *Hist. de l'Académie de Paris* 1705.

**) *Ars conjectandi* Lib. IV. Cap. IV. pag. 226.

ganz besonders betonen. Denn es scheint nicht berechtigt zu sein, wenn Laplace in seiner Notice historique sur le calcul des probabilités bei der Erwähnung der Verdienste Daniel Bernoulli's sagt:*) «On doit «surtout placer au nombre de ces idées originales la considération di- «recte des possibilités des événemens tirées des événemens observés. «Jacques Bernoulli et Moivre supposaient ces possibilités connues; et «ils cherchaient la probabilité que le résultat des expériences à faire «approchera de plus en plus de les représenter.»

Nicht Daniel, wie aus dem Citat hervorgehen möchte, sondern Jakob Bernoulli ist der Begründer der Theorie von der Erfahrungswahrscheinlichkeit. Er hat auch den ersten analytischen Ausdruck dafür gegeben.**) Wenn in einer Urne sich weisse und schwarze Kugeln befinden, deren Zahlenverhältniss aber unbekannt ist, so wird, wenn man in einer sehr grossen Anzahl von Versuchen a weisse und b schwarze herausgezogen hat, die Wahrscheinlichkeit für das Ziehen einer weissen ausgedrückt durch $\frac{a}{a+b}$.

Auch über die Wahrscheinlichkeit der Ursachen hat Jakob Bernoulli zuerst Untersuchungen angestellt.***) Gewiss hatte er noch tiefere analytische Studien über die Wahrscheinlichkeit a posteriori vorgesehen, wahrscheinlich auch praktische Versuche auf socialem Gebiete, aber leider wurde Bernoulli† viel zu früh, schon mit 51 Jahren, der Wissenschaft durch den Tod entrissen und ein halbes Jahrhundert ging dahin, bis er richtig verstanden wurde, bis Daniel Bernoulli, sein Neffe, praktisch und Bayes theoretisch seine Untersuchungen über die Erfahrungswahrscheinlichkeit weiter führten.

*) Essai philosophique p. 214. Théorie analyt. des prob. introd. p. CXLVIII.

**) Ars conj. Lib. IV. Cap. IV.

***) id. Lib. IV. Cap. III.

†) Jakob Bernoulli I., in Basel als Sohn des Rathsherrn Nikolaus Bernoulli am 27. XII. 1654 geboren, studirte in seiner Vaterstadt Theologie und daneben fleissig Mathematik. Nach seinem theologischen Examen (1676) bereiste er die Schweiz, Holland, England und Frankreich, widmete sich dann nach seiner Rückkehr als Privatmann ganz der Mathematik und wurde im Jahre 1687 zum Professor der Mathematik an der Universität Basel ernannt, welche Stellung er bis zu seinem Tode am 16. VIII. 1695 innehaltete. Mit seinem Bruder Johannes I. und seinem Neffen Daniel gehört Jakob Bernoulli I. zu den berühmtesten der Bernoulli.

III.

8. Abraham de Moivre*) war der erste, der dem Theorem Bernoulli's gebührende Aufmerksamkeit schenkte und dasselbe in geschickter Weise zu fördern verstand.

Indem aber Moivre dasselbe nicht wie Bernoulli vom Standpunkt der Wahrscheinlichkeit a posteriori aus anfasste, sondern als Untersuchungsobjekt für sich, trachtete er darnach, für den Fall, in welchem die einfache Wahrscheinlichkeit eines Ereignisses E als bekannt und constant gleich p , diejenige des entgegengesetzten E' gleich q vorausgesetzt wird, einen bestimmten Werth zu suchen für die Wahrscheinlichkeit, dass in einer grossen Anzahl von μ Versuchen das Ereigniss E in einer solchen Anzahl m von Malen eintreffe, die zwischen den Grenzen $p - 1$ und $p + 1$ liegt, d. h. einen bestimmten Werth zu geben für den Bernoulli'schen Summenausdruck

$$\sum_{m=\mu-p-1}^{\mu-p+1} \frac{\mu!}{m! n!} p^m q^n$$

Zwei Schwierigkeiten mussten ihm dabei entgegentreten, die Auffindung eines allgemeinen, numerisch leicht zu berechnenden Ausdrucks für den Binomialcoefficienten resp. für die Facultät und die Summation der Terme einer binomischen Entwicklung innerhalb gewisser Grenzen. Unsere weitere historische Untersuchung wird daher in der Folge eine Periode der Geschichte der Summationsformeln in sich einbeziehen müssen.

Moivre hat die Hauptresultate seiner Untersuchungen über das Bernoulli'sche Theorem niedergelegt in einem grössern Abschnitt seiner *Doctrine of chances***), betitelt: *A Method of approximating the Sum*

*) Abraham de Moivre (geb. 26. V. 1667 in Vitry, Champagne, gest. 27. XI. 1754 in London), protestantischer Refugié, durch den Widerruf des Edikts von Nantes durch Louis XIV. 1785 genöthigt, in London ein Asyl zu suchen, erwarb sich dort lange Zeit durch Privatstunden kümmerlich sein Brot. Später genoss er die Protektion Newtons und wurde 1697 Mitglied der Royal Society. Neben seinen Hauptwerken, die in der Arbeit citirt sind, schrieb er: *A new method for valuing of annuities upon lives*. Der nach ihm benannte Lehrsatz findet sich auf der ersten Seite seiner *Miscell. anal.* Der grosse Newton soll in den letzten Jahren seines Lebens zu denjenigen, welche ihm mathematische Fragen vorlegten, gesagt haben: «Go to Mr. Moivre, he knows these things better than I do.» Ein ehrenderes Zeugniss konnte Moivre wohl nicht gegeben werden.

**) P. 235 ff. Uns lag die 2. Auflage (London 1738) vor; zum ersten Male erschien das bedeutende Werk im Jahre 1718 unter dem Titel: *De mensura sortis*.

of the Terms of the Binomial $(a + b)^n$ expanded into a series from whence are deduced some practical Rules to estimate the Degree of Assent which is to be given to Experiments*). Die analytischen Erläuterungen zu den Resultaten dieser Untersuchung gibt Moivre zerstreut in seinem andern Buche *Miscellanea analytica de serieb. et quadrat.*, und es mag nicht ganz ohne Werth sein, hier eine zusammenhängende Darstellung derselben zu geben.

9. In die oben erwähnte Abhandlung einleitend, erwähnt Moivre die Schwierigkeit der Summation von Gliedern einer binomischen Entwicklung und er hat für seine Zeit vollkommen Recht. Selbst die grossen Mathematiker Jakob und Nikolaus Bernoulli hätten eigentlich nicht eine Summe von solchen Gliedern gegeben, sondern nur weite Grenzen gezeigt, in welchen sich eine gewisse Summe derselben bewegen könne. Moivre sagt dann weiter: Es sind mehr als 12 Jahre verflossen**) seit ich gefunden habe, dass wenn man das Binom $(1 + 1)^n$ entwickelt, der mittlere Term zur Summe aller Terme — zu 2^n — ein Verhältniss hat, das gleich ist

$$\frac{2 A (n - 1)^n}{n^n \sqrt{n-1}}, \text{ worin}$$

$$\text{Log } A = \frac{1}{12} - \frac{1}{360} + \frac{1}{1260} - \frac{1}{1680} \pm \dots \dots$$

Für $n = \infty$, folgt

$$\text{Log } \frac{(n - 1)^n}{n^n} = \text{Log } \left(1 - \frac{1}{n}\right)^n = -1 \text{ und}$$

setzt man

*) Die Abhandlung findet sich in lateinischer Uebersetzung auch als Anhang in der *Miscell. analyt.* Sie war schon vor der *Doctrine of chances* im Druck erschienen, jedoch nicht veröffentlicht worden. Es geht dies aus folgender interessanten Bemerkung hervor, die Moivre im Zusatz von Problem 87, wo er über die Schwierigkeiten des Problems sich ausspricht, macht: «I take the liberty to say, that this is the hardest Problem, that can be proposed on the subject of chance, for which reason I have reserved it for the last, but I hope to be forgiven if my solution is not fitted to the capacity of all Readers; however I shall derive from it some Conclusions that may be of use to everybody: in order thereto here translate a Paper of mine, which was printed Nov. 12. 1733, and communicated to some Friends, but never yet made public, reserving to myself the right of enlarging my own Thoughts, as occasion shall require.»

**) Es war also ums Jahr 1720.

$$\log B = -1 + \frac{1}{12} - \frac{1}{360} + \frac{1}{1260} - \frac{1}{1680} \pm \dots$$

so wird das obige Verhältniss für $n = \infty$ gleich $\frac{2B}{\sqrt{n}}$, oder wenn man die Gleichung von $\log B$ mit -1 multiplizirt, wird es gleich $\frac{2}{B\sqrt{n}}$.

Ueber den Werth von B äussert sich Moivre auf folgende Weise:
 « When I first began that inquiry, I contented myself to determine at large the Value of B , which was done by the addition of some Terms of the above-written Series; but as I perceiv'd that it converged but slowly*), and seeing at the same time that what I had done answered my purpose tolerably well, I desisted from proceeding farther, till my worthy and learned Friend Mr. James Stirling, who had applied himself after me to that inquiry, found that the Quantity B did denote the square-root of the Circumference of a Circle whose Radius is Unity, so that if that Circumference be called c the Ratio of the middle Term to the sum of all the Terms will be expressed by $\frac{2}{\sqrt{nc}}$. »

10. Ueber diesen eleganten Ausdruck für sein gesuchtes Verhältniss war Moivre hocherfreut. Wie er aber zum Ausdruck

$$\frac{2A(n-1)^n}{n^n \sqrt{n-1}}$$

gekommen ist, darüber finden wir Auskunft in *Miscellanea analytica de serieb. et quadrat.* Lib. VI. Cap. II.: *De regressu et Serie data ad Summam.* Hier führt Moivre aus: Der Coefficient c des mittleren Gliedes im Binom $(1+1)^n$ ist, wenn man $\frac{n}{2} = m$ setzt

$$c = \frac{(m+1)(m+2)(m+3)\dots\dots\dots 2m}{(m-1)(m-2)(m-3)\dots\dots\dots 2 \cdot 1 \cdot m},$$

und es wird

*) Moivre war in dieser Convergenzfrage im Irrthum. Denn die Reihe $1 - \frac{1}{12} + \frac{1}{360} - \frac{1}{1260} \pm \dots \dots \dots$ ist gleich der divergenten Reihe $1 - \frac{B(1)}{1 \cdot 2} + \frac{B(2)}{3 \cdot 4} - \frac{B(3)}{5 \cdot 6} \pm \dots \dots \dots$, wenn $B(1), B(2), \dots \dots \dots$ die Bernoulli'schen Zahlen bedeuten.

$$\begin{aligned}
 \log c &= 2 \left(\frac{1}{m} + \frac{1}{3m^3} + \frac{1}{5m^5} + \frac{1}{7m^7} + \dots \text{in inf.} \right) \\
 &+ 2 \left(\frac{2}{m} + \frac{8}{3m^3} + \frac{32}{5m^5} + \frac{128}{7m^7} + \dots \text{in inf.} \right) \\
 &+ 2 \left(\frac{3}{m} + \frac{27}{3m^3} + \frac{243}{5m^5} + \frac{2187}{7m^7} + \dots \text{in inf.} \right) \\
 &+ \dots \dots \dots \dots \dots \dots \dots \dots \\
 &\vdots \\
 &+ 2 \left(\frac{m-1}{m} + \frac{(m-1)^3}{3m^3} + \frac{(m-1)^5}{5m^5} + \frac{(m-1)^7}{7m^7} + \dots \text{in inf.} \right)^*) \\
 &+ \log 2.
 \end{aligned}$$

Nimmt man aus diesen $m - 1$ Logarithmenreihen die Colonnen zusammen, so wird

$$\begin{aligned}
 \log c &= \frac{2}{m} \left(1 + 2 + 3 + 4 + \dots + m - 1 \right) \\
 &+ \frac{2}{3m^3} \left(1 + 2^3 + 3^3 + 4^3 + \dots + (m-1)^3 \right) \\
 &+ \frac{2}{5m^5} \left(1 + 2^5 + 3^5 + 4^5 + \dots + (m-1)^5 \right) \\
 &+ \dots \dots \dots \dots \dots \dots \dots \dots \\
 &\vdots \\
 &\text{in inf.} \quad + \log 2.
 \end{aligned}$$

Bezeichne ich nun die Reihen nach einander mit I, II, III, ... und setze $m - 1 = s$, so wird nach den Tafeln Jakob Bernoulli's:

$$\begin{aligned}
 \text{I} &= \frac{s^2 + s}{m} \\
 \text{II} &= \frac{\frac{s^4}{2} + s^3 + \frac{s^2}{2}}{3m^3} \\
 \text{III} &= \frac{\frac{s^6}{3} + s^5 + \frac{5s^4}{6} + \frac{s^2}{6}}{5m^5} \\
 \text{IV} &= \frac{\frac{s^8}{4} + s^7 + \frac{7s^6}{6} - \frac{7s^4}{12} + \frac{s^2}{6}}{7m^7} \\
 \text{V} &= \frac{\frac{s^{10}}{5} + s^9 + \frac{3s^8}{2} - \frac{7s^6}{2} + s^4 - \frac{3s^2}{10}}{9m^9} \\
 &\text{in inf.}
 \end{aligned}$$

*) Die Convergenz der letzten Reihen ist allerdings sehr gering.

Die Colonnen dieser Summen wieder in Reihen zusammengefasst ergibt als erste, wenn $\frac{s}{m} = x$ gesetzt wird:

$$s \left(x + \frac{x^3}{2 \cdot 3} + \frac{x^5}{3 \cdot 5} + \frac{x^7}{4 \cdot 7} + \dots \text{ in inf.} \right)$$

$$= s \left(\frac{2x}{1 \cdot 2} + \frac{2x^3}{3 \cdot 4} + \frac{2x^5}{5 \cdot 6} + \frac{2x^7}{7 \cdot 8} + \dots \text{ in inf.} \right).$$

Entwickle ich $\log \frac{1+x}{1-x} = v$ in die logarithmische Reihe und multiplizire beiderseits mit x^*), so kommt

$$\frac{2x \dot{x}}{1} + \frac{2x^3 \dot{x}}{3} + \frac{2x^5 \dot{x}}{5} + \dots \text{ in inf.} = vx,$$

und nimmt man auf beiden Seiten die Fluenten (d. h. integriert man), so hat man

$$\frac{2x^2}{1 \cdot 2} + \frac{2x^4}{3 \cdot 4} + \frac{2x^6}{5 \cdot 6} + \dots \text{ in inf.}$$

$$= x \log \frac{1+x}{1-x} - \log \frac{1}{1-x^2}.$$

Auf beiden Seiten mit $\frac{s}{x}$ multiplizirt, erhält, weil $\frac{s}{m} = x$, die Gleichung die Form

$$\frac{s^2}{m} + \frac{s^4}{2 \cdot 3 m^3} + \frac{s^6}{3 \cdot 5 m^5} + \dots \text{ in inf.}$$

$$= m x \log \frac{1+x}{1-x} - m \log \frac{1}{1-x^2},$$

oder weil $s = m - 1$, und $x = \frac{s}{m} = \frac{m-1}{m}$, so erhält man leicht für die Summe der

1. Colonne $= (2m-1) \log (2m-1) - 2m \log m$.

Die 2. Colonne besteht aus folgender Reihe

$$\frac{s}{m} + \frac{s^3}{3m^3} + \frac{s^5}{5m^5} + \dots \text{ in inf.} = \frac{1}{2} \log \frac{m+s}{m-s}$$

$$= \frac{1}{2} \log (2m-1).$$

Die Summe beider Colonnen wird daher gleich

$$(2m - \frac{1}{2}) \log (2m-1) - 2m \log m.$$

*) Moivre bezeichnetet, wie es bei den englischen Mathematikern im vorigen Jahrhundert nach dem Vorgange Newtons üblich war, das Differential dx mit x .

Nimmt man dazu noch den beiseite gesetzten Log 2, so wird das Summenaggregat für Log c in erster Näherung

$$(2m - \frac{1}{2}) \log(2m - 1) - 2m \log m + \log 2.$$

Subtrahirt man hievon den $\log 2^{2m} = 2m \log 2$, so bleibt

$$(2m - \frac{1}{2}) \log(2m - 1) - 2m \log 2m + \log 2,$$

und dieser Ausdruck wird, weil $2m = n$, wenn man zugleich zur Exponentialfunktion übergeht, zu

$$\frac{2(n-1)^{n-\frac{1}{2}}}{n^n}$$

und dies ist der angenäherte Wert des Verhältnisses des mittleren Coeffizienten des mittleren Gliedes in der Entwicklung von $(1+1)^n$ zur Summe aller Glieder.

Im gegebenen Ausdruck sind aber nur die beiden ersten Colonnen des logarithmischen Summenaggregates für Log c berücksichtigt, während es deren unendlich viele gibt. Die 3. Colonne konstituiert die geometrische Progression:

$$\frac{s^2}{6m^3} + \frac{s^4}{6m^5} + \frac{s^6}{6m^7} + \dots \text{ in inf.} = \frac{1}{6m} \cdot \frac{(m-1)^2}{2m-1}.$$

Die 4. Colonne gibt die recurrente Reihe:

$$\frac{s^2}{180m^5} \left\{ 6 + \frac{15s^2}{m^2} + \frac{28s^4}{m^4} + \frac{45s^6}{m^6} + \frac{66s^8}{m^8} + \dots \text{ in inf.} \right\}.$$

deren Beziehungsscala 3, -3, +1, ist und deren Summe gefunden wird, als

$$- \frac{(4m^4 + 2m^3 + 3m^2 - 4m + 1)(m-1)^2}{180m^3(2m-1)^3}$$

Indem ich bemerkte — fährt Moivre fort — dass diese Reihen obwohl durchaus summirbar, doch sehr verwickelt werden, brachte ich sie auf den Fall des Unendlichen. Wird $m = \infty$, so ist der Werth der 3. Colonne gleich $\frac{1}{12}$, von der 4. gleich $-\frac{1}{360}$, wie man aus obigen Formeln leicht finden kann. Für die 5. und 6. Colonne habe ich die Werthe $+\frac{1}{1260}$, resp. $-\frac{1}{1680}$ gefunden.

Wird der Numerus der Logarithmenreihe Moivre's

$$\frac{1}{12} - \frac{1}{360} + \frac{1}{1260} - \frac{1}{1680} \pm \dots \text{ in inf.}$$

mit A bezeichnet, so wird das Verhältniss des mittleren Gliedes im entwickelten Binom $(1 + 1)^n$ — wenn $n =$ sehr gross vorausgesetzt wird — zur Summe aller Glieder ausgedrückt durch $\frac{2 A (n - 1)^{n - \frac{1}{2}}}{n^n}$, wie ihn Moivre in der Doctrine of chances gegeben hat.

11. Als Resultat eines Versuches, die Constante $2A$ durch Addition der 4 ersten Terme seiner logarithmischen Reihe zu bestimmen, fand Moivre die Zahl 2,168. Dieser Versuch darf hier aber um so eher übergangen werden, weil Moivre durch Bemerkungen*) von Seiten Stirlings selber zur Einsicht von der Unzulänglichkeit seiner Methode gelangte und in dem *Miscellaneis Analyticis Supplementum* die Untersuchung neu begann. Er bemerkt dort einleitend :

«Attamen post receptam Stirlingii Epistolam, cum mihi aliquid
«vacui temporis suppeteret, constitueram totum illud denuo excutere,
«atque initium sumere ab isto Problemate de inveniendis Summis Lo-
«garithmorum ab unitate incipientium; ecce autem gradus quibus ad
«meam solutionem adductus sum, quam ideo trado quod modus solutionis
«quo utor sit longe diversus ab eo quem Stirlingius adhibuit, quo fiet
«ut suspicio a me aberit me voluisse actum agere».

Moivre sucht also hier direct wie Stirling — auf den wir noch zurückkommen werden — die Summe der Logarithmen der natürlichen Zahlen und nicht mehr wie früher das Verhältniss des Coeffizienten vom mittleren Gliede zur Summe aller Glieder. Er geht aus vom Product:

$$\frac{m}{m-1} \cdot \frac{m}{m-2} \cdot \frac{m}{m-3} \cdots \cdots \frac{m}{m-m+2} \cdot \frac{m}{m-m+1},$$

und entwickelt die Logarithmen der einzelnen Factoren in folgender Darstellung:

$$\log \frac{m}{m-1} = -\frac{1}{m} + \frac{1}{2m^2} - \frac{1}{3m^3} + \dots \text{ in inf.}$$

$$\text{Log } \frac{m}{m-2} = -\frac{2}{m} + \frac{4}{2m^2} + \frac{8}{3m^3} + \dots \text{ in inf.}$$

$$\text{Log } \frac{m}{m-3} = \frac{3}{m} + \frac{9}{2m^2} + \frac{27}{3m^3} + \dots \dots \dots \text{ in inf.}$$

$$\text{Log } \frac{m}{m-m+1} = \frac{m-1}{m} + \frac{(m-1)^2}{2m^2} + \frac{(m-1)^3}{3m^3} \dots \text{ in inf.}$$

*) In einem Briefe an Moivre vom 17. Juni 1729. Vergl. Misc. analyt. Lib. VII.

Analog wie beim Coeffizientenproblem stellt nun Moivre die Colonnen als Zeilen zusammen, deren er unendlich viele erhält. Die Zähler jeder dieser Zeilen stellen eine Reihe dar von der Form:

$$1^n + 2^n + 3^n + \dots + (m-1)^n.$$

Moivre summirt diese Potenzreihen, dividirt jede Summe durch den zugehörigen Nenner und erhält so, indem er $m - 1 = 1$ setzt, folgende neue Reihen als Summen der obigen Colonnen:

Moivre nimmt wieder die Colonnen als Reihen zusammen, dividiert die erste durch m , die dritte durch $\frac{A}{2m}$, die vierte durch

$\frac{B}{3 \cdot 4 \text{ m}^3}$, die fünfte durch $\frac{C}{5 \cdot 6 \text{ m}^5}$, , worin

$$A = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

$$B = \frac{1}{2} - \frac{1}{5} - \frac{4}{2} A = - \frac{1}{30}$$

$$C = \frac{1}{2} - \frac{1}{7} - \frac{6}{2} A - \frac{6 \cdot 5 \cdot 4}{2 \cdot 3 \cdot 4} B = \frac{1}{42}$$

$$D = \frac{1}{2} - \frac{1}{9} - \frac{8}{2} A - \frac{8 \cdot 7 \cdot 6}{2 \cdot 3 \cdot 4} B - \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} C = -\frac{1}{30}$$

also die Bernoullischen Zahlen bedeuten, und erhält die folgenden neuen Beihen:

$$\begin{aligned}
 \frac{1^2}{2m} + \frac{1^3}{6m^3} + \frac{1^4}{12m^4} + \frac{1^5}{20m^5} + \dots &\text{ in inf.} \\
 \frac{1}{2m} + \frac{1^2}{4m^2} + \frac{1^3}{6m^3} + \frac{1^4}{8m^4} + \dots &\text{ in inf.} \\
 \frac{1}{m} + \frac{1^2}{m^2} + \frac{1^3}{m^3} + \frac{1^4}{m^4} + \dots &\text{ in inf.} \\
 \frac{3}{m} + \frac{61^2}{m^2} + \frac{101^3}{m^3} + \frac{151^4}{m^4} + \dots &\text{ in inf.} \\
 \frac{5}{m} + \frac{151^2}{m^2} + \frac{351^3}{m^3} + \frac{701^4}{m^4} + \dots &\text{ in inf.} \\
 \frac{7}{m} + \frac{281^2}{m^2} + \frac{841^3}{m^3} + \frac{2101^4}{m^4} + \dots &\text{ in inf.} \\
 \dots &\text{ in inf.}
 \end{aligned}$$

Die ersten beiden dieser Reihen sind logarithmische. Die Summe der ersten findet man auf folgende Weise:

$$\begin{aligned}
 \text{Sei } v = \log \frac{1}{1-x}, \text{ so ist, wenn man entwickelt,} \\
 xdx + \frac{1}{2}x^2dx + \frac{1}{3}x^3dx + \frac{1}{4}x^4dx + \dots \text{ in inf.} = vdx.
 \end{aligned}$$

Integriert man, so wird

$$\begin{aligned}
 \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{12}x^4 + \frac{1}{20}x^5 + \dots \text{ in inf.} &= vx + x - v \\
 &= x \log \frac{1}{1-x} + x - \log \frac{1}{1-x}.
 \end{aligned}$$

Oder für $x = \frac{1}{m}$ gesetzt, so erhält man weil $1 = m - 1$ als Summe der ersten Reihe

$$\begin{aligned}
 \frac{1}{2} \frac{(m-1)^2}{m^2} + \frac{1}{6} \frac{(m-1)^3}{m^3} + \frac{1}{12} \frac{(m-1)^4}{m^4} + \dots \text{ in inf.} \\
 &= \frac{m-1 - \log m}{m}.
 \end{aligned}$$

Multipliziert man noch mit m (wodurch man früher dividirt hat), so wird der gesuchte Werth der Reihe

$$\frac{1}{2} \frac{(m-1)^2}{m} + \frac{1}{6} \frac{(m-1)^3}{m^2} + \dots \text{ in inf.} = m - 1 - \log m.$$

Die zweite Reihe hat folgende Summe, wie unmittelbar folgt:

$$\frac{1}{2} \frac{m-1}{m} + \frac{1}{4} \frac{(m-1)^2}{m^2} + \frac{1}{6} \frac{(m-1)^3}{m^3} + \dots \text{ in inf.} = \frac{1}{2} \log m.$$

Die beiden ersten Reihen konstituiren also die Summe

$$m - 1 - \frac{1}{2} \log m.$$

Die Summen der übrigen Reihen lassen sich rational und zwar auf folgende Weise ausdrücken (indem man zugleich wieder mit jenen Factoren, mit welchen die Reihen multiplizirt wurden, dividirt):

$$\text{Summe der 3. Reihe} = \frac{A}{2m} (m-1) = \frac{1}{12} - \frac{1}{12m}.$$

$$\text{“ “ “ 4. “} = \frac{B}{3 \cdot 4 m^3} (m^3 - 1) = - \frac{1}{360} + \frac{1}{360m^3} \text{*)}.$$

Analog wird:

$$\text{“ “ “ 5. “} = \frac{1}{1260} - \frac{1}{1260m^5}.$$

$$\text{“ “ “ 6. “} = - \frac{1}{1680} + \frac{1}{1680m^7}.$$

.....

Es wird somit:

$$\begin{aligned} \log \left[\frac{m}{m-1} \cdot \frac{m}{m-2} \cdot \frac{m}{m-3} \cdots \frac{m}{2} \cdot \frac{m}{1} \right] &= m - 1 - \frac{1}{2} \log m \\ &- \frac{1}{12m} + \frac{1}{360m^3} - \frac{1}{1260m^5} + \frac{1}{1680m^7} \mp \dots \text{ in inf.} \\ &+ \frac{1}{12} - \frac{1}{360} + \frac{1}{1260} - \frac{1}{1680} \pm \dots \text{ in inf.}, \end{aligned}$$

oder

$$\begin{aligned} \log \left[1 \cdot 2 \cdot 3 \cdots (m-1) \right] &= \left(m - \frac{1}{2} \right) \log m \\ &- m + \frac{1}{12m} - \frac{1}{360m^3} + \frac{1}{1260m^5} - \frac{1}{1680m^7} \pm \dots \text{ in inf.} \\ &+ 1 - \frac{1}{12} + \frac{1}{360} - \frac{1}{1260} + \frac{1}{1680} \mp \dots \text{ in inf.} \end{aligned}$$

Dies ist die *Moivre'sche Reihe für Log I(m)*. Fügt man derselben noch $\log m$ bei, bezeichnet man die Summe der Constanten mit C und führt in der Reihe

$$\frac{1}{12m} - \frac{1}{360m^3} + \frac{1}{1260m^5} \mp \dots \text{ in inf.}$$

*) Es ist $3x + 6x^2 + 15x^3 + 21x^4 + \dots \text{ in inf.} = \frac{1}{1-x^3} - 1$
also

$$3 \frac{m-1}{m} + 6 \frac{(m-1)^2}{m^2} + 15 \frac{(m-1)^3}{m^3} + \dots \text{ in inf.} = m^3 - 1.$$

die Bernoulli'schen Zahlen ein, so geht dieselbe über in die folgende Summationsformel:

$$\sum_{m=1}^{\infty} \log m = c + \left(m + \frac{1}{2} \right) \log m - m + \frac{B(1)}{1 \cdot 2 \cdot m} - \frac{B(2)}{3 \cdot 4 m^3} + \frac{B(3)}{5 \cdot 6 m^5} - \frac{B(4)}{7 \cdot 8 m^7} \pm \dots,$$

welche sich auch leicht aus der allgemeinen Summationsformel, die Euler, wie später gezeigt werden soll, in den Inst. Calc. Diff. Part. II, C. V. aufgestellt hat, ergibt, nämlich aus der Formel:

$$\sum z = \int z dx + \frac{1}{2} z + \frac{B(1)}{2!} \frac{dz}{dx} - \frac{B(2)}{4!} \frac{d^3 z}{dx^3} + \frac{B(3)}{6!} \frac{d^5 z}{dx^5} \pm \dots \dots \dots,$$

wenn man für $z = \log m$ setzt.

Es verdient daher hier hervorgehoben zu werden, dass Moivre zuerst, wenn auch empirisch, diese Summationsformel angewendet hat.

12. Im Weitern gibt Moivre in den Supplementa auch eine besser convergirende Reihe für die Constante, d. h. für die Reihe

$$1 - \frac{1}{12} + \frac{1}{360} - \frac{1}{1260} + \frac{1}{1680} \mp \dots \dots \dots \text{ in inf.},$$

welche nach seiner Ansicht «satis commode convergit in principio «post terminos quinque primos convergentiam amittit, quam tamen «postea recuperat». Indem er $m - 1 = 9$ setzt, erhält er nach seiner Formel:

$$1 - \frac{1}{12} + \frac{1}{360} - \frac{1}{1280} + \frac{1}{1680} \mp \dots \text{ in inf.} = \log 5040.72 \\ - 9\frac{1}{2} \log 10 + 10 - \frac{1}{12 \cdot 10} + \frac{1}{360 \cdot 10^3} - \frac{1}{1260 \cdot 10^5} \pm \dots \text{ in inf.}$$

Den cyklometrischen Charakter der Constanten hat aber Moivre nicht von sich aus erkannt; denn er war sehr erstaunt darüber, als ihm Stirling in einem Schreiben) vom 19. Juni 1729 mittheilte, dass der Werth der Constanten $\sqrt{2\pi}$ betrage. «Nemo est profecto «qui post visam hanc superioris problematis solutionem fateri recuset «eam esse usquequaque mirabilem: sed nihil in ea fortasse mirabilius «videbitur quam qua arte Quadratura Circuli potuerit in eam induci», sagt Moivre über Stirlings Lösung. Er spricht dort auch die Vermuthung aus, Stirling habe sein Resultat mit Hülfe der Formel von*

*) Veröffentlicht in *Miscellanea analyt. Cap. VII.*

Wallis gefunden. Er selbst habe die Sache *deshalb* nicht weiter verfolgt, weil ihm die Lösung nur ein *Mittel zur Berechnung von Wahrscheinlichkeiten* sei. «Adde quod cum non ideo susceptum fuisset ut «propter se solveretur, sed ut juvaret solutionem alterius cuiusdam «problematis quod pulcherrimum judicaveram, mihi videbar in iis «quae feceram aliquo jure posse acquiescere.»*)

IV.

13. Es erscheint hier geboten, Moivre und seine *Doctrine of chances* für einen Augenblick zu verlassen, um in Stirlings mathematischem Werke: *Methodus differentialis sive Tractatus de summatione et interpolatione serierum infinitarum***) nach der Bestimmung der Constanten $\frac{1}{2} \log 2\pi$ zu sehen.

Stirling findet dieselbe zuerst bei der Berechnung des Verhältnisses, welches der Coeffizient des mittleren Gliedes einer binomischen Entwicklung zur Summe aller Coeffizienten hat. Die Priorität der Lösung dieses Problems erkennt er aber ausdrücklich Moivre zu, wenn er am Schluße der Vorrede zu seinem Buche sagt: «Problema de «invenienda Uncia media in permagna dignitate binomii solutam erat «a Moivraeo ante aliquot annos quam ego idem attingeram: Nec pro- «babile est quod in hunc usque diem de eodem cogitasse, in sugges- «sisset Spectatissimus Vir, D. Alex. Cuming***) se plurunum suspicari «an idem solvi posset per Methodum Differentialem Newtoni.»

Stirling gibt zwei verschiedene Methoden zur Lösung des Coeffizientenproblems, wovon die eine, die auf Interpolation mit Hilfe der Differenzenrechnung beruht†), hier nicht berücksichtigt werden soll, weil dort die Bestimmung der Constanten auf numerischer Berechnung††) beruht.

Die andere Methode†††) ist nach ihm folgende:

Sei gegeben die Reihe:

$$1, 2, \frac{8}{3}, \frac{16}{5}, \frac{128}{35}, \frac{256}{63}, \dots$$

*) *Miscellaneis analyticis Supplementum* p. 3.

**) London 1730.

***) Alex. Cuming darf in der Geschichte der Wahrscheinlichkeitsrechnung nicht unerwähnt bleiben. Aus Bemerkungen, die Moivre in *Miscell. analyt. Cap. V* macht, geht hervor, dass derselbe auch ihm manche Anregungen gegeben hat. Ueber das Leben Cumings habe ich nichts in Erfahrung gebracht.

†) Dargestellt im Propos. XXII, Ex. I p. 116 ff.

††) Vergl. Note 2 im Anhang.

†††) *Method. diff. Propos. XXIII.*

deren Glieder das gesuchte Verhältniss für resp. die 0., 1., 2., 3., Potenz reciprok darstellen, so handelt es sich um die Interpolation des allgemeinen Gliedes der Reihe :

$$1, \frac{2A}{1}, \frac{4B}{3}, \frac{6C}{5}, \frac{8D}{7}, \dots,$$

wenn mit A, B, C, D, allgemein unsere Reihe ausgedrückt wird.

Sei nun T irgend ein Glied dieser Reihe, so wird das folgende Glied, wenn wir einer Variablen n die Werthe 0, 2, 4, 6, geben, gleich sein

$$\begin{aligned} T_1 &= \frac{n+2}{n+1} T, \text{ oder} \\ T_1^2 &= \frac{n^2 + 4n + 4}{n^2 + 2n + 1} T^2 \text{ und} \\ 2T_1^2 + (n+2)(T^2 - T_1^2) - \frac{T_1^2}{n+2} &= 0 \quad . \quad \alpha) \end{aligned}$$

Man setze nun

$$T^2 = an + \frac{bn}{n+2} + \frac{cn}{(n+2)(n+4)} + \frac{dn}{(n+2)(n+4)(n+6)} + \dots, \beta)$$

worin a, b, c, d, noch zu bestimmende Constante bedeuten ; diese Reihe in andere Form gebracht, wird zu

$$T^2 = an + b + \frac{c - 2b}{n+2} + \frac{d - 4c}{(n+2)(n+4)} + \dots$$

Analog:

$$T_1^2 = a(n+2) + b + \frac{c - 2b}{(n+4)} + \frac{d - 4c}{(n+4)(n+6)} + \dots$$

hieraus

$$\begin{aligned} (n+2) \{ T^2 - T_1^2 \} &= -2a(n+2) + \frac{2c - 4b}{n+4} \\ &+ \frac{4d - 16c}{(n+4)(n+6)} + \dots \end{aligned}$$

Substituirt man neben diesem Werthe noch diejenigen für $2T_1^2$ und für $-\frac{T_1^2}{n+2}$ in die Gleichung α , so kommt:

$$2b - a + \frac{4c - 9b}{n+4} + \frac{6d - 25c}{(n+4)(n+6)} + \frac{8c - 49d}{(n+4)(n+6)(n+8)} + \dots = 0.$$

Aus dieser Gleichung ergeben sich für die Coeffizienten die folgenden Bedingungsgleichungen :

$$\begin{array}{l}
 2b - a = 0 \\
 4c - 9b = 0 \\
 6d - 25c = 0 \\
 8e - 49d = 0
 \end{array}
 \quad \left| \begin{array}{l}
 \text{also: } b = \frac{a}{2} \\
 c = \frac{9b}{4} = \frac{9a}{2 \cdot 4} \\
 d = \frac{25c}{6} = \frac{9 \cdot 25a}{2 \cdot 4 \cdot 6} \\
 e = \frac{49d}{8} = \frac{9 \cdot 25 \cdot 49a}{2 \cdot 4 \cdot 6 \cdot 8}
 \end{array} \right.$$

Werden diese Werthe in die Gleichung β substituirt, so ergibt sich:

$$\begin{aligned}
 T^2 &= a(n + \frac{n}{2(n+2)} + \frac{9n}{2 \cdot 4(n+2)(n+4)} \\
 &\quad + \frac{9 \cdot 25n}{2 \cdot 4 \cdot 6(n+2)(n+4)(n+6)} + \dots) \\
 &= an(1 + \frac{1}{2(n+2)} + \frac{9}{2 \cdot 4(n+2)(n+4)} \\
 &\quad + \frac{9 \cdot 25}{2 \cdot 4 \cdot 6(n+2)(n+4)(n+6)} + \dots)
 \end{aligned}$$

Den Coeffizienten a bestimmt nun Stirling durch folgende Ueberlegung: Je grösser n , desto wahrer wird die Gleichung

$$T^2 = an.$$

Setzt man nun in dieser Gleichung für n der Reihe nach seine Werthe 0, 2, 4, und die entsprechenden für T^2 , so erhält man eine Reihe von Näherungsgleichungen für a , nämlich:

$$a = 2, = 2 \cdot \frac{8}{9}, = 2 \cdot \frac{8}{9} \cdot \frac{24}{25}, = 2 \cdot \frac{8}{9} \cdot \frac{24}{25} \cdot \frac{48}{49}, = \dots$$

Daher ist der Werth von a gleich dem ins Unendliche fortgesetzten Produkt

$$2 \cdot \frac{8}{9} \cdot \frac{24}{25} \cdot \frac{48}{49} \cdot \frac{80}{81} \cdot \frac{120}{121} \dots,$$

dessen Werth aber nach der Formel von Wallis gleich $\frac{\pi}{2}$ ist.

Es resultirt somit für T folgender Werth:

$$T = \sqrt{\frac{\pi n}{2} \left[1 + \frac{1}{2(n+2)} + \frac{9}{2 \cdot 4(n+2)(n+4)} + \frac{9 \cdot 25}{2 \cdot 4 \cdot 6(n+2)(n+4)(n+6)} + \dots \right]}.$$

Oder es ist nach Annahme, wenn man mit M den Coeffizienten des mittleren Termes der binomischen Entwicklung bezeichnet, mit S die Summe aller Coeffizienten:

$$M:S = 1 : \sqrt{\frac{n\pi}{2} \left[1 + \frac{1}{2(n+2)} + \frac{9}{2 \cdot 4(n+2)(n+4)} + \dots \right]}.$$

Stirling gelangt daher zu folgendem Satze:

Der Exponent des Binoms, wenn gerade, sei n , wenn ungerade $n - 1$; dann wird sich der mittlere Coeffizient zur Summe aller Coeffizienten verhalten, wie die Einheit zur mittleren Proportionale zwischen dem halben Kreisumfang und der einen oder anderen von folgenden Reihen:

$$n + \frac{A}{2(n+2)} + \frac{9B}{4(n+4)} + \frac{25C}{6(n+6)} + \dots,$$

$$n + 1 - \frac{A}{2(n-3)} - \frac{9B}{4(n-5)} - \frac{25C}{6(n-7)} - \dots,$$

wenn man allgemein die Reihen nach Newton'scher Bezeichnung mit

$$A \pm B \pm C \pm D \pm \dots \dots \dots \text{ darstellt.}$$

Ueber den Gebrauch der Formel spricht sich Stirling dahin aus, es genüge, wenn $n ==$ sehr gross werde, zu setzen

$$T^2 = \frac{\pi}{2} \left(n + \frac{1}{2} \right), \text{ oder}$$

$$T = \sqrt{\frac{\pi}{2} \left(n + \frac{1}{2} \right)}.$$

Es ist also

$$\begin{aligned} M:S &= 1 : \sqrt{\frac{\pi}{2} \left(n + \frac{1}{2} \right)} = \frac{1}{\sqrt{\frac{\pi}{2} \left(n + \frac{1}{2} \right)}} \\ &= \frac{2}{\sqrt{2\pi n}} \text{ (für } \lim n = \text{ sehr gross).} \end{aligned}$$

Das Stirling'sche Resultat beim Coeffizientenproblem ist somit demjenigen Moivre's genau gleich; denn Moivre hat für das Verhältniss des mittleren Gliedes zur Summe aller im entwickelten Binom $(1+1)^n$, für $n =$ gerade, den nämlichen Ausdruck, jedoch ohne cyklometrische Darstellung der Constanten, gefunden.

14. Wie Moivre, so musste auch Stirling durch das Coeffizientenproblem darauf kommen, einen numerisch leicht zu berechnenden Summenausdruck für $\log \Gamma(x)$ resp. für $\Gamma(x)$ zu suchen. Er behandelt dazu folgende Aufgabe*): Es sei die Summe beliebig vieler Logarithmen zu finden, deren Numeri in arithmetischer Progression fortschreiten.

*) Loc. cit. Propos. XXVIII. p. 135.

Es mögen $x + n, x + 3n, x + 5n, x + 7n, \dots, z - 3n, z - n$, beliebig viele Zahlen in arithmetischer Progression bezeichnen, die letzte sei $z - n$. Es seien ferner $\log z$ und $\log x$ die Tafellogarithmen der Zahlen z und x , und a sei gleich dem Modul, d. h. gleich dem reciproken Werth des Log. nat. von 10. Dann wird die Summe der Logarithmen der vorliegenden Reihe gleich sein der Differenz zwischen den beiden folgenden Reihen:

$$\begin{aligned} \frac{z \log z}{2n} &= \frac{az}{2n} - \frac{an}{12z} + \frac{7an^3}{360z^2} - \frac{31an^5}{1260z^5} + \frac{127an^7}{1680z^7} + \dots \\ &\quad \vdots \dots \dots \text{ in inf.} \\ \frac{x \log x}{2n} &= \frac{ax}{2n} - \frac{an}{12x} + \frac{7an^3}{360x^2} - \frac{31an^5}{1260x^5} + \frac{127an^7}{1680x^7} \\ &\quad \vdots \dots \dots \text{ in inf.} \end{aligned}$$

Diese Reihen setzen sich so ins Unendliche fort:

Man setze

$$\begin{aligned} -\frac{1}{3.4} &= A, \quad -\frac{1}{5.8} = A + 3B, \\ -\frac{1}{7.12} &= A + 10B + 5C, \\ = \frac{1}{9.16} &= A + 21B + 35C + 7D, \\ &\quad \dots \dots \dots \end{aligned}$$

Die Zahlen, die in den verschiedenen Werthen von $A, B, C, D \dots$ multiplizirt werden, sind die ersten, dritten, fünften, \dots Coeffizienten der ungeraden Potenzen des Binoms. Dann wird der Coeffizient des dritten Terms $= \frac{1}{12} = A$, der des vierten $B = \frac{7}{360}$, der des fünften $C = -\frac{31}{1260}$ und so fort.

Beweis. Es werde die Variable z um ihre Abnahme (constante Differenz) $2n$ vermindert, d. h. man substituire $z - 2n$ für z in die Reihe

$$\frac{z \log z}{2n} - \frac{az}{2n} - \frac{an}{12z} + \frac{an^3}{360z^2} - \frac{31an^5}{1260z^5} \pm \dots \text{ in inf.}$$

Man subtrahire die neue Reihe von der vorigen, so wird sich, nachdem man durch Division die Terme auf die nämliche Form gebracht, als Rest ergeben:

$$\log z - \frac{an}{z} - \frac{an^2}{2z^2} - \frac{an^3}{3z^3} + \dots \dots \dots \text{ in inf.}$$

d. h. der Logarithmus der Zahl $z - n$.

So ist allgemein die Abnahme zweier successiven Werthe der Reihe gleich dem Logarithmus von $z - n$, der allgemein jeden beliebigen der Logarithmen bedeuten kann, welche zu summiren sind. Die Reihe wird also die Summe der vorgegebenen Logarithmen sein, wenn von ihr die andere Reihe subtrahirt wird. Denn die Summen der Reihen sind wie diejenigen der Flächen zuweilen zu corrigiren, damit sie richtig werden (Constante).

In Exemp. II, al. 2 geht Stirling alsdann so weiter: *Will man die Summen von beliebig vielen Logarithmen der natürlichen Zahlenreihe 1, 2, 3, $z - n$ haben, so ist $n = \frac{1}{2}$, und es werden*

3 oder 4 Glieder der Reihe

$$z \log z - az - \frac{a}{24z} + \frac{7a}{2880z^3} + \dots \dots \dots ,$$

zu denen man den halben Logarithmus des Kreisumfanges, dessen Radius die Einheit ist, d. h. 0.39908 zu addiren hat, die gewünschte Summe geben und zwar mit um so weniger Mühe, je mehr Logarithmen zu summiren sind (Convergenz).

15. Dies die Stirling'sche Darstellung seiner nach ihm benannten Reihe. Stirling findet also zunächst*), zwar ohne ein Verfahren anzugeben, für

$$\begin{aligned} \log(x + n) + \log(x + 3n) + \log(x + 5n) + \dots \dots \dots \\ + \log(z - 3n) + \log(z - n) = \end{aligned}$$

die Differenz der beiden Reihen von natürlichen Logarithmen:

$$\begin{aligned} \frac{z \log z}{2n} - \frac{z}{2n} - \frac{n}{12z} + \frac{7n^3}{360z^3} - \frac{31n^5}{1260z^7} \pm \dots \dots \text{ in inf.} \\ \frac{x \log x}{2n} - \frac{x}{2n} - \frac{n}{12x} + \frac{7n^3}{360x^3} - \frac{31n^5}{1260x^7} \pm \dots \dots \text{ in inf.} \end{aligned}$$

Um die Richtigkeit seines Satzes zu beweisen, erhärtet dann Stirling denselben für den Spezialfall $x = z - 2n$. Handelt es sich aber um den Logarithmus des Produktes 1. 2. 3, so wird, da $n = \frac{1}{2}$ und $x = \frac{1}{2}$ ist,

*) Jedenfalls durch Entwicklung der einzelnen Logarithmen.

$$\begin{aligned} \text{Log} \left(1 \cdot 2 \cdot 3 \dots \left(z - \frac{1}{2} \right) \right) &= z \text{ Log } z - z - \frac{1}{24z} \\ &\quad + \frac{7}{2880 z^3} \mp \dots \dots \\ - \left\{ \frac{1}{2} \text{ Log } \frac{1}{2} - \frac{1}{2} - \frac{1}{12} + \frac{7}{360} - \frac{31}{1260} \pm \dots \dots \text{ in inf.} \right\}. \end{aligned}$$

Stirling gibt als Resultat nur die erste dieser Reihen mit der Bemerkung, man habe dazu noch $\frac{1}{2} \text{ Log } 2\pi$ zu addiren.

Wie oft in seinem Buche, gibt Stirling auch hier nur das Resultat, ohne zugleich den Weg zu weisen, auf welchem er dazu gelangt ist, was das Studium desselben sehr erschwert. Es entzieht sich daher einer sicheren Beurtheilung, wie Stirling die Constante bestimmt hat. Eine numerische Berechnung scheint mir ausgeschlossen zu sein, gerechtfertigter aber erscheint die Vermuthung, dass er auch hier wie beim Coeffizientenproblem die Formel von Wallis angewendet hat und am meisten Wahrscheinlichkeit besitzt wohl die Annahme, derselbe habe in diesem Falle die Constante durch Vergleichung mit dem Resultate des Coeffizientenproblems gefunden.)*

16. Stirling gab**) schon, was hier Erwähnung verdient, das Euler-sche Integral 1. Art., nämlich :

$$B(r+z, p-r) = \int_0^1 x^{r+z-1} (1-x)^{p-r-1} dx$$

und benutzte dasselbe zur Interpolation z. B. der Reihe

$$a, \frac{ra}{p}, \frac{(r+1)b}{p+1}, \frac{(r+2)c}{p+2}, \dots,$$

indem er für das allgemeine Glied T der Reihe fand :

*) Wenn $n = \text{gerade}$, so gilt nach Stirling und Moivre in der Entwicklung von $(1+1)^n$

$$\frac{n!}{\left(\frac{n}{2}\right)! \left(\frac{n}{2}\right)! 2^n} = \sqrt{\frac{2}{2n\pi}}$$

woraus man, wenn für die Fakultäten der Stirling'sche Näherungswert substituirt wird, die Constante bestimmen kann.

**) L. c. Propos. XXIV, p. 126.

$$\frac{T}{a} = \frac{\int_0^1 x^{r+z-1} (1-x)^{p-r-1} dx}{\int_0^1 x^{r-1} (1-x)^{p-r-1} dx},$$

also $T = \frac{a \Gamma(r+z) \Gamma(p)}{\Gamma(r) \Gamma(p+z)}.$

Ebenso fand er als intermediäres Glied T zwischen dem ersten und zweiten der folgenden Reihe

$$1, \frac{1}{2}, \frac{1 \cdot 3}{2 \cdot 4}, \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}, \dots,$$

$$T = \frac{\int_0^1 \frac{1}{\sqrt{1-x}} dx}{\int_0^1 \frac{1}{\sqrt{x-x^2}} dx} = \frac{2}{\pi}$$

und *gewiss ist nicht daran zu zweifeln, Stirling war nahe daran, die Näherungswerte für den Binomialcoffizienten und die Fakultät auf analogem Wege zu suchen, wie es Laplace später gethan**), nämlich mit Hülfe der sogenannten Euler'schen Integrale.

17. Es scheint, dass Stirling über sein Verfahren, die Constante zu bestimmen, auch in keiner andern Publikation**) Auskunft gegeben hat, denn Moivre schrieb noch 1738:

«But altho it be not necessary to know what relation the number «B may have to the Circumference of the Circle, provided its value «be attained, either by pursuing the Logarithmic Series before men-«tioned, or any other way; yet I own with pleasure that this discovery, «besides that it has saved trouble, has spread a singular Elegancy on «the Solution.»

Bezeichnet man in der Stirling'schen Reihe $z - \frac{1}{2}$ mit m und führt die Bernoullischen Zahlen ein, so ergibt sich folgende Summationsformel:

*) Vergl. Note 4 im Anhang.

**) Doctrine of chances, p. 236.

$$\begin{aligned}\text{Log}(1 \cdot 2 \cdot 3 \dots m) &= \frac{1}{2} \text{Log} 2\pi + \left(m + \frac{1}{2}\right) \text{Log} \left(m + \frac{1}{2}\right) \\ &- \left(m + \frac{1}{2}\right) - \frac{(2-1) B(1)}{1 \cdot 2 \cdot 2(m + \frac{1}{2})} + \frac{(2^3 - 1) B(2)}{3 \cdot 4 \cdot 2^3 (m + \frac{1}{2})^3} - \dots,\end{aligned}$$

welche eine von der Moivre'schen etwas abweichende Form hat. Aus beiden Formeln aber ergibt sich für $\lim m = \infty$, wenn man zur Exponentialfunktion übergeht:

$$\lim_{m \rightarrow \infty} m! = m^m e^{-m} \sqrt{2\pi m} ,$$

welche Formel auch die *Stirling'sche* genannt wird.

Es ist unstreitig das Verdienst des mit mathematischem Scharf-sinn ausserordentlich begabten Stirling), die Constante $\sqrt{2\pi}$ bestimmt zu haben. Berücksichtigt man aber, dass Moivre zuerst das Coeffizientenproblem gestellt und gelöst hat und dass derselbe auch die andere Aufgabe, die sich aus jenem ergeben musste, die Summe der Logarithmen der natürlichen Zahlen zu suchen, unabhängig und fast gleichzeitig mit Stirling ebenfalls gelöst hat, vergisst man nicht, dass Moivre diese Formel zuerst in der Wahrscheinlichkeitsrechnung, für welche ihr grosse Bedeutung zukommt, praktisch verwendet hat, so muss man sagen, dass dessen Name mit der Formel in ebenso verdienstvollem Sinne verbunden ist, wie derjenige Stirlings.*

Die Ursprungsgeschichte der Stirling'schen Formel aber ist ganz besonders geeignet, zu zeigen, wie befruchtend eine angewandte mathematische Disziplin auf die reine Mathematik wirken kann.

V.

18. Nachdem hiemit die Untersuchungen Moivres und Stirlings über das *Coeffizientenproblem* und über die *Summe von Log I(x)* sowohl unter sich wie auch in ihrem gegenseitigen Verhältniss gewürdigt sind, kehren wir wieder zu Moivres Abhandlung über das *Bernoullische Theorem* in dessen «Doctrine of chances» zurück.

*) James Stirling, geb. 1696 in St. Ninians, Grafschaft Stirling, Schottland, gest. 5. Dez. 1770 in Leadhiks, studirte in Oxford Mathematik, bewarb sich als Agent einer schottischen Bergbaugesellschaft erfolglos um eine Professur. Er wurde schon 1729 Mitglied der Royal Society. Sein Hauptwerk, *Methodus differentialis*, erlebte 3 Auflagen (1730, 1753, 1764), war aber schon 1718 unvollständig in den Philos. Transact. erschienen.

Als zweiten analytischen Fundamentalsatz gibt Moivre folgenden*):
Der Logarithmus des Verhältnisses, welches der Coeffizient des mittleren Termes einer binomischen Entwicklung von sehr hoher Potenz n in Bezug auf den Coeffizienten irgend eines um das Intervall l von ihm entfernten Termes hat, wird in erster Näherung durch folgende Grösse ausgedrückt:

$$\left(m + 1 - \frac{1}{2}\right) \log(m + 1 - 1) + \left(m - 1 + \frac{1}{2}\right) \log(m + 1 - 1) \\ - 2m \log m + \log \frac{m + 1}{m},$$

vorausgesetzt, dass $m = \frac{n}{2}$ gesetzt wird.

Sein Lösungsverfahren für dieses Resultat ist ein analoges wie beim Coeffizientenproblem, geht also aus von logarithmischen Reihen (v. Miscell. analyt. p. 128 ff.) und es braucht daher hier nicht wiederholt zu werden.

Moivre zieht dann weiter aus dem angeführten Satze die folgenden hier skizzirten Schlüsse in Form von Zusätzen.

Z u s a t z 1. Wenn $m = \frac{n}{2}$ eine unendliche Grösse bedeutet, so ist *der Logarithmus des Verhältnisses, welches ein Term (immer in der Entwicklung $(1 + 1)^n$) der vom mittleren Term um das Intervall l entfernt ist, zum letzteren hat, gleich $-\frac{2l^2}{n}$.*

Z u s a t z 2. Die Zahl, deren hyperbolischer Logarithmus $-\frac{2l^2}{m}$, ist gleich der Reihe

$$1 - \frac{2l^2}{n} + \frac{4l^4}{2n^2} - \frac{8l^6}{6n^3} + \frac{16l^8}{24n^4} - \frac{32l^{10}}{120n^5} \pm \dots \text{ in inf.}$$

woraus folgt**), dass *die Summe der Terme vom grössten an bis und mit jenem, der um l Glieder entfernt ist, gleich ist:*

$$\frac{2}{\sqrt{2n\pi}} \left\{ 1 - \frac{2l^3}{1 \cdot 3n} + \frac{4l^5}{2 \cdot 5n^2} - \frac{8l^7}{6 \cdot 7n^3} + \frac{16l^9}{24 \cdot 9n^4} \mp \dots \text{ in inf.} \right\}$$

Setzt man nun $1 = s\sqrt{n}$, alsdann wird die Summe:

$$\frac{2}{\sqrt{2\pi}} \left\{ s - \frac{2s^3}{1 \cdot 3} + \frac{4s^5}{2 \cdot 5} - \frac{8s^7}{6 \cdot 7} + \frac{16s^9}{24 \cdot 9} \mp \dots \text{ in inf.} \right\}$$

*) Loc. cit. p. 236.

**) Moivre gibt keine weitere Begründung dieser Folgerung.

und für $s = \frac{1}{2}$, entsteht die Reihe:

$$\frac{2}{\sqrt{2\pi}} \left\{ \frac{1}{2} - \frac{1}{3 \cdot 4} + \frac{1}{2 \cdot 5 \cdot 8} - \frac{1}{6 \cdot 7 \cdot 16} + \frac{1}{24 \cdot 9 \cdot 32} - \frac{1}{120 \cdot 11 \cdot 64} + \dots \right\}.$$

Durch Addition von 7 oder 8 Gliedern dieser ziemlich gut convergirenden Reihe erhält man nach einfacher logarithmischer Rechnung als Verhältniss der Summe der 1 Terme zwischen dem mittleren und dem um 1 entfernten in der Entwicklung von $(1 + 1)^n$ zur Summe aller Terme die Zahl 0,341344.

Z u s a t z 3. Hat ein Ereigniss dieselbe einfache und constante Wahrscheinlichkeit auf Eintreffen wie auf Nichteintreffen, so wird, wie aus den Prinzipien der Wahrscheinlichkeitsrechnung hervorgeht, die Wahrscheinlichkeit, dass das Ereigniss bei n Versuchen höchstens $\frac{n}{2} + 1$ und wenigstens $\frac{n}{2} - 1$ Mal eintreffe, ausgedrückt durch $\frac{S}{2^n}$, wenn S die Summe aller Terme in der Entwicklung von $(1 + 1)^n$, genommen zwischen den Gliedern, die um 1 Terme links und rechts vom mittleren abstehen (die äussersten inbegriffen), bedeutet. Die Wahrscheinlichkeit also, dass *ein Ereigniss unter gleichen Verhältnissen in einer solchen Zahl von Malen eintrifft, die zwischen* $\frac{n}{2} \pm \frac{1}{2} \sqrt{n}$ *liegt, ist daher gegeben durch das Doppelte der Zahl, die im Zusatz 2 gefunden wurde, durch 0,682688 und die Wahrscheinlichkeit des Gegentheils, dass die Eintreffenzahl ausserhalb dieser Grenzen fällt, ist somit 0,317312.*

Z u s a t z 4. Weil es aber unausführbar ist, eine unendliche Zahl von Experimenten anzustellen, so können wir den vorhergehenden Schluss auch auf grosse endliche Zahlen anwenden (folgt ein Beispiel für $n = 3600$).

Z u s a t z 5. Wir können daher als fundamentale Maxime hinstellen: *Das Verhältniss, welches in der Entwicklung des Binoms von hoher Potenz die Summe der Glieder, welche vom mittleren Term aus nach beiden Seiten um ein Intervall von $\frac{1}{2}\sqrt{n}$ Gliedern liegen, zur Summe der ganzen Entwicklung hat, wird ausgedrückt durch die Zahl 0,682688 oder nahezu $\frac{28}{41}$; hiebei ist aber nicht nöthig, dass*

$n = \infty$ sei; sogar für $n = 100$ liefert die Regel noch ein erträgliches Resultat, wie ich durch Versuche bestätigt finde. Noch ist zu bemerken, dass $\frac{1}{2}\sqrt{n}$, im Verhältniss bezogen auf n , um so kleiner wird, je mehr n wächst; wächst also die Zahl der Beobachtungen, so werden die Grenzen im Verhältniss zu n immer enger während die Wahrscheinlichkeit dieselbe bleibt.

Zusatz 6. Wenn $l = \sqrt{n}$ gesetzt wird, so konvergiert die Reihe in Coll. 2 weniger gut als für $l = \frac{1}{2}\sqrt{n}$, und für eine erträgliche Annäherung sind daher viel mehr Terme zu addiren. In diesem Falle gebrauche ich die mechanische Quadratur, die von Sir Isaac Newton erfunden, von Mr. Cotes*), Mr. James Stirling und mir, vielleicht noch von anderen weiter ausgebildet worden ist. Sie besteht in der Bestimmung der Fläche einer Curve, wenn man von ihr eine gewisse Anzahl von Ordinaten A, B, C, D, \dots kennt, die sich in gleichen Intervallen folgen, wobei auch gilt, dass, je kürzer die Intervalle genommen werden, desto genauer das Resultat wird. Im vorliegenden Falle beschränke ich mich auf 4 Ordinaten, die mit A, B, C, D bezeichnet sein mögen. Wenn nun der Abstand der ersten von der letzten gleich 1 ist, so wird die Fläche gleich

$\frac{(A + D) + 3(B + C)}{8} \cdot l$ sein**). Setzen wir nun die Distanzen

gleich 0, $\frac{1}{6}\sqrt{n}$, $\frac{2}{6}\sqrt{n}$, $\frac{3}{6}\sqrt{n}$, $\frac{4}{6}\sqrt{n}$, $\frac{5}{6}\sqrt{n}$, und \sqrt{n} , verwenden für unsern Fall die 4 letzten: $\frac{3}{6}\sqrt{n}$, $\frac{4}{6}\sqrt{n}$, $\frac{5}{6}\sqrt{n}$, $\frac{6}{6}\sqrt{n}$, nehmen alsdann die Quadrate dieser Ausdrücke, verdoppeln jeden, dividiren durch n und geben jedem das Zeichen minus, so haben wir die Größen:

$-\frac{1}{2}$, $-\frac{8}{9}$, $-\frac{25}{18}$, -2 , welche die hyperbolischen Logarithmen der

*) Cotes Roger (10. VII. 1682 — 5. VI. 1716), Professor der Astronomie und Physik in Cambridge, war der Verfasser der *Harmonia mensurarum* (Cambridge 1722), welche den bekannten Cotesischen Lehrsatz enthält.

**) Moivre leitet diese Formel (*Miscell. analyt. lib. VII c. II: «De Methodo differentiarum»*) aus der Newton'schen Interpolationsformel ab, nämlich aus:

$$u_n = u + \binom{u}{1} \Delta u + \binom{n}{2} \Delta^2 u + \dots + \binom{u}{n} \Delta^n u,$$

worin u_n das allgemeine Glied, u das Anfangsglied und $\Delta u, \Delta^2 u, \Delta^3 u \dots$ die Anfangsglieder der ersten, zweiten, dritten \dots Differenzreihen sind.

Zahlen 0,60653, 0,41111, 0,24935, 0,13534 sind, die unsere 4 Ordinaten darstellen. Weil nun $1 = \frac{1}{2}\sqrt{n}$ ist, so ergibt sich nach der Formel für unsere Fläche $0,170203\sqrt{n}$. Das Doppelte hievon multiplizirt mit $\sqrt{2n\pi}^2$ ergibt die Zahl 0,27160, und diese zu 0,682688 (Zusatz 7) addirt gibt 0,95428, welches die Wahrscheinlichkeit ist, dass bei n Versuchen das Ereigniss weder mehr als $\frac{n}{2} + \sqrt{n}$, noch weniger als $\frac{n}{2} - \sqrt{n}$ eintritt.

Zusatz 7. Auf demselben Wege kann man finden, wie gross die Wahrscheinlichkeit sei, dass die Zahl des Eintreffens zwischen andern Grenzen liege, z. B. zwischen $\frac{n}{2} \pm \frac{3}{2}\sqrt{n}$. Hiefür würde sich die Zahl 0,99874 finden lassen.

Bei allen Beispielen spielt \sqrt{n} die Rolle eines Modulus für die Schätzung der Grenzen und der Wahrscheinlichkeiten.

Zusatz 8. Ist die einfache und constante Wahrscheinlichkeit eines Ereignisses nicht gleich der entgegengesetzten, bildet die Zahl der günstigen zu den ungünstigen Fällen das Verhältniss $\frac{a}{b}$, so lässt sich die Wahrscheinlichkeit dafür, dass das Ereigniss in n Versuchen eine solche Zahl von Malen eintreffe, die zwischen $\frac{an}{a+b} \pm l$ liegt, ausdrücken durch $\frac{S}{(a+b)^n}$, wo S die Summe aller Glieder in der binomischen Entwicklung von $(a+b)^n$ bedeutet, die links und rechts im Intervall von 1 Gliedern (die äussersten inbegriffen) vom grössten Gliede abstehen. Das Verhältniss, welches bei einer sehr hohen Potenz des Binoms $(a+b)^n$ das grösste Glied der Entwicklung zur Summe aller übrigen Glieder hat, wird ausgedrückt durch den Bruch $\frac{a+b}{\sqrt{abn\pi}}^*$.

Zusatz 9. Der Logarithmus des Verhältnisses, welches ein Term in der binomischen Entwicklung, der um das Intervall von 1 Termen vom grössten absteht, zu diesem hat, ist gleich $-\frac{a+b}{2abn} l^2$.

*) Meines Wissens gibt Moivre nirgends eine analytische Herleitung weder von dieser Formel, noch jener im Zusatz 9. Die Lösung ergibt sich jedoch analog wie jene bei Voraussetzung gleicher entgegengesetzter Wahrscheinlichkeiten.

Z u s a t z 10. Ist die Wahrscheinlichkeit eines Ereignisses auf Eintreffen verschieden von derjenigen auf Nichteintreffen, so werden die Probleme, die Summation der Terme in der Entwicklung von $(a + b)^n$ betreffend, mit derselben Leichtigkeit und Methode aufgelöst wie diejenigen, wo die entgegengesetzten Wahrscheinlichkeiten dieselben sind.

Aus dem Gesagten folgt, dass der Zufall die Ereignisse, die natürlichen Institutionen gemäss eintreten, sehr wenig in ihrem Eintreffen stört. Wird z. B. ein rundes Metallstück, dessen Seiten fein polirt sind und verschiedene Farben, z. B. schwarz und weiss zeigen, aufgeworfen, so wird mit der Vermehrung der Würfe das Verhältniss der erhaltenen Schwarz und Weiss sich immer mehr der Gleichheit nähern und es ist schon bei 3600 Versuchen die Wahrscheinlichkeit dafür, dass die Erscheinungszahl der einen oder andern Farbe zwischen 1770 und 1830 liege annähernd $\frac{2}{3}$; in diesem Falle macht also die Abweichung von der perfekten Gleichheit nur $\frac{1}{120}$ der gesammten Versuchszahl aus und mit derselben Wahrscheinlichkeit wäre die Abweichung bei 10,000 Versuchen nur $\frac{1}{2000}$ aller Erscheinungen. Mit der Erweiterung der Grenzen aber würde die Wahrscheinlichkeit für das Eintreffen einer der Farben in einer Anzahl von Malen, die in diesen Grenzen liegt, immer wachsen und schliesslich zur Gewissheit werden. Diese Ausdehnung der Grenzen aber, und das ist nicht zu vergessen, ist bei Vermehrung der Beobachtungen im Vergleich zum Wachsthum der Versuchszahl nicht so beträchtlich, diese wächst direct, jene mit der Quadratwurzel.

Schliesslich müsste also bei unendlich vielen Versuchen mit Gewissheit eine Gleichheit unter der Zahl der Erscheinungen von Schwarz und Weiss eintreten.

Die nämliche Betrachtung liesse sich auch durchführen für den Fall, in welchem die entgegengesetzten Wahrscheinlichkeiten ungleiche sind.

Abraham de Moivre schliesst seine werthvolle Abhandlung mit einer Ueberlegung, die an Jakob Bernoulli's kühne Schlusskonsequenzen erinnert: «And thus in all cases it will be found, that altho Chance «produces Irregularities, still the Odds will be infinitely great, that in «process of Time, those Irregularities will bear no proportion to the «recurrency of that Order which naturally results from original Design.»*)

*) Doctrine of chances, 2. ed. p. 243.

19. J. Todhunter hält Moivre neben Laplace für den grössten Analytiker in der Wahrscheinlichkeitsrechnung, wenn er sagt*): «*It will not be doubted that the theorie of Probability owns more to him than to any other mathematician, with the sole exception of Laplace.*»

Pflichtet man diesem Urtheil ohne Einschränkung bei, so muss insbesondere noch hervorgehoben werden, dass kein Mathematiker um die analytische Darstellung des Bernoulli'schen Theorems grössere Verdienste hat als Moivre. Nicht von vornehmerein von einer so hohen philosophischen Warte ausschauend wie Jakob Bernoulli und sich demnach nicht weiter über die Wahrscheinlichkeit a posteriori verbreitend, schenkte Moivre der mathematischen Analyse des Problems sein Hauptinteresse, und erfolgreich hat er die heutigen Methoden und Resultate der analytischen Darstellung desselben im Prinzip gegeben.

Es gelang Moivre nicht nur, mit Stirlings Hülfe einen leicht zu berechnenden Ausdruck für die Fakultät zu finden, sondern er hat auch schon als Summe von Termen einer binomischen Entwicklung innerhalb gewisser Grenzen den Laplace'schen Integralausdruck gegeben.

Denn :

Bezeichnet M das Mittelglied der Entwicklung von $(1 + 1)^n$, M_1 das um ein Intervall von 1 Gliedern entfernte Glied, so wird nach Moivre (v. Zusatz 2)**):

$$1) \frac{M_1}{M} = e^{-\frac{2l^2}{n}} = 1 - \frac{2l^2}{n} + \frac{4l^2}{2n^2} + \frac{8l^6}{6n^3} \pm \dots \dots \dots \text{ in inf.}$$

Wie nun Moivre die Summe der Terme zwischen M und M_1 gefunden, sagt er nirgends; es lässt sich aber annehmen, dass er die Ausdrücke der linken Seite der folgenden Gleichung in Exponentialreihen entwickelt und summirt hat :

$$\begin{aligned} \frac{M_1}{M} + \frac{M_{1-1}}{M} + \dots + \frac{M_1}{M} + \frac{M}{M} &= e^{-\frac{2l^2}{n}} + e^{-\frac{2(1-1)^2}{n}} \\ &+ e^{-\frac{2(1-2)^2}{n}} + \dots + e^{-\frac{2}{n}} + e^0, \end{aligned}$$

woraus sich ergibt

$$M_1 + M_{1-1} + \dots + M_1 + M = M. \text{ [Summe der Exponentialr.]}$$

Moivre erhält dann, indem er noch durch die Summe der ganzen Entwicklung dividirt den Ausdruck :

$$\frac{2}{\sqrt{2n\pi}} \left\{ 1 - \frac{2l^3}{1 \cdot 3n} + \frac{4l^5}{2 \cdot 5n^2} - \frac{8l^7}{6 \cdot 7n^3} + \frac{16l^9}{24 \cdot 9n^4} \pm \dots \text{ in inf.} \right\}.$$

*) Todhunter, History of the Prob. p. 193.

**) Bei den folgenden Hinweisen auf Zusätze sind immer diejenigen in Moivre's Abhandlung gemeint.

Man wird sich nun leicht überzeugen, dass der Ausdruck in Parenthese weniger jene [Summe der Exponentialgrössen] darstellt, sondern genau das unbestimmte Integral der Reihe in Gleichung 1), d. h., Moivre nimmt für das Verhältniss der Summe der Terme von M bis M_1 (inclus. die äussersten) zur Summe aller Terme das bestimmte Integral, welches man gewöhnlich als Laplace'sches bezeichnet

$$\frac{2}{\sqrt{2n\pi}} \int_0^1 e^{-\frac{2x}{n}} dx.$$

Um die Wahrscheinlichkeit zu erhalten, dass bei n Versuchen die Zahl der günstigen Beobachtungen sich innerhalb der Grenzen $\frac{n}{2} \pm 1$ liege, verdoppelt Moivre den Werth jenes Integrals (Zusatz 3) und erhält somit allgemein für die bezeichnete Wahrscheinlichkeit :

$$W = \frac{4}{\sqrt{2n\pi}} \int_0^1 e^{-\frac{2x^2}{n}} dx,$$

oder im besondern Fall, wenn $1 = \frac{1}{2} \sqrt{n}$ gesetzt wird

$$W = 0,682688.$$

Für den Fall, in welchem die entgegengesetzten einfachen Wahrscheinlichkeiten ungleich sind, würde Moivre nach Zusatz 9 für W erhalten:

$$W = \frac{2(a+b)}{\sqrt{2abn\pi}} \int_0^1 e^{-\frac{a+b}{2abn}x^2} dx.$$

Dieser Deduction haften zwei Ungenauigkeiten an. Zunächst wird das *mittlere grösste Glied zweimal gezählt*. Dieser Fehler compensirt sich zwar bei gleichen einfachen und entgegengesetzten Wahrscheinlichkeiten, wenn die Versuchszahl n als ungerade Zahl vorausgesetzt wird, in welchem Falle dann 2 Mittelglieder vorhanden sind.

Im Weiteren benutzt Moivre offenbar die *Summationsformel* :

$$\sum_{x=0}^{x=1} \varphi(x) = \int_0^1 \varphi(x) dx.$$

Wie aber im nächsten Abschnitt gezeigt werden soll, hat MacLaurin zuerst gefunden und Euler es auf andere Weise bestätigt, dass für eine stetige, nach endlichen Incrementen fortschreitende Funktion in erster Näherung die Formel gilt (wenn die Variable sehr gross wird):

$$\sum_{x=0}^{l+1} \varphi(x) = \int_0^{l+1} \varphi(x) dx - \frac{1}{2} \left[\varphi(x) \right]_0^{l+1}.$$

Darnach würde, bei gleichen entgegengesetzten, constanten Wahrscheinlichkeiten, wenn man die *Moivre'sche Funktion* $e^{-\frac{2x^2}{n}}$, die Stetigkeit besitzt und für $x = 0$ ein Maximum liefert, statt $\varphi(x)$ setzt und unter der Voraussetzung, dass die Versuchszahl n eine gerade ist,

$$W = \frac{2}{\sqrt{2n\pi}} \left[2 \int_0^{l+1} e^{-\frac{2x^2}{n}} dx - e^{-\frac{2(l+1)^2}{n}} \right]$$

und im andern Falle, wenn die Versuchszahl ungerade,

$$W = \frac{2}{\sqrt{2n\pi}} \left[2 \int_0^{l+1} e^{-\frac{2x^2}{n}} dx - e^{-\frac{2(l+1)^2}{n}} + 1 \right].$$

Ungeachtet dieser Ungenauigkeiten, die sich wohl begreifen lassen, bleibt *Moivre der Schöpfer des Laplace'schen Integrals* und hat überhaupt das Verdienst, die Infinitesimalrechnung zuerst in der Wahrscheinlichkeitstheorie fruchttragend verwerthet zu haben (z. B. auch beim Coeffizientenproblem). Ferner hat Moivre *zum ersten Mal* eine *Wahrscheinlichkeitscurve* angenommen, einzelne Flächentheile derselben durch mechanische Quadratur bestimmt (Zusatz 6) und deren Wendepunkte angegeben*). Interessant ist auch, dass Moivre im Falle von gleichen entgegengesetzten einfachen Wahrscheinlichkeiten die *Wendepunktsordinate* resp. den Term für $l = \frac{1}{2} \sqrt{n}$ (Zusatz 2) als *Fehlergrenze wählt*. Diese spielt heute bekanntlich in der Fehlertheorie**) eine wichtige Rolle, weil sich aus ihr ein charakteristischer Fehler, welcher der Wurzel aus dem mittleren Fehlerquadrat entspricht, ergibt.

Was die Analysis aus den Moivre'schen Wahrscheinlichkeitsstudien für sich gewonnen, braucht nach alledem nicht mehr weiter ausgeführt zu werden; dagegen möchten wir schliesslich noch der *logischen Klarheit und Uebersichtlichkeit* in Moivres analytischen Entwicklungen, die man bei Stirling oft vermisst und worin Moivre vielleicht der Lehrer der Meister in dieser Hinsicht — *Euler* und *Lagrange* — geworden ist, lobend gedenken.

*) Vergl. Note 3 im Anhang.

**) S. Hagen, *Grundzüge der Wahrscheinlichkeitsrechnung*, p. 73 ff.

VI.

20. Zeigte sich im letzten Abschnitt die Unzulänglichkeit des Moivre'schen Verfahrens für die Ueberführung einer nach endlichen Incrementen fortschreitenden Summe zum Integral, so geht hinwieder aus den Abschnitten III und IV hervor, dass die Summationsformeln von Moivre und Stirling zur angenäherten Bestimmung eines Werthes für $\log I(x+1)$ mehr empirischer Natur waren und daher der Allgemeingültigkeit ermangelten. Aber bis um die Mitte des vorigen Jahrhunderts hatte sich die Analysis schon bedeutend entwickelt, und es musste sich in der Reihentheorie selbst das Bedürfniss nach allgemeinen Summationsformeln geltend machen.

MacLaurin*) war der erste, der auf Grund der von Newton begründeten mechanischen Quadratur eine allgemeinere Summationsformel für Reihen mit endlichen Differenzen aufstellte. Er betrachtet**) eine parabolische Curve von der Gleichung:

$$y = A + Bz + Cz^2 + Dz^3 + \dots,$$

oder wenn a die Anfangsordinate bezeichnet,

$$y = a + \frac{zda}{dz} + \frac{z^2d^2a}{2! dz^2} + \frac{z^3d^3a}{3! dz^3} + \dots$$

MacLaurin setzt nun $dz = 1$ und bezeichnet mit A, B, C, D, \dots die Flächen, deren gemeinsame Basis gleich dz und deren Ordinaten respective $y, dy, d^2y, d^3y \dots$ sind und findet für

$$A = a + \frac{da}{2!} + \frac{d^2a}{3!} + \frac{d^3a}{4!} + \dots$$

dann wird

$$a = A - \frac{da}{2!} - \frac{d^2a}{3!} - \frac{d^3a}{4!} - \dots$$

Werden nun auf analoge Weise $da, d^2a, d^3a, d^4a \dots$ bestimmt, wie z. B.

$$da = B - \frac{d^2a}{2!} - \frac{d^3a}{3!} - \frac{d^4a}{4!} - \dots$$

\dots

so ergibt sich schliesslich durch Substitution:

$$a = A - \frac{B}{2} + \frac{C}{12} - \frac{E}{720} + \frac{G}{30240} \mp \dots$$

oder allgemein:

*) Colin MacLaurin, geboren zu Killnoddan in Schottland im Jahre 1698, war Professor der Mathematik zu Aberdeen und Edinburgh. Er starb 1746.

**) Treatise of Fluxions (Edinburgh 1742) art. 830. a. fs.

$$a = A - KB + LC - MD + NE \mp \dots \dots \dots$$

worin die Coeffzienten $K, L, M, N \dots \dots \dots$, wenn man $k = \frac{1}{2!}$, $l = \frac{1}{3!}$, $m = \frac{1}{4!}$, $\dots \dots \dots$ setzt, nach folgendem Gesetze fortschreiten:

$$K = k = \frac{1}{2}$$

$$L = kK - l = \frac{1}{12}$$

$$M = kL - lK + m = 0$$

$$N = kM - lL + mK - n = - \frac{1}{720}$$

.....,

so dass also die Coeffzienten der Flächen $D, F, H \dots \dots$ verschwinden. Nun ist A gleich dem Integral von ydz , B dasjenige von $dy dz$, C von $d^2y dz$, $\dots \dots$, alle Integrale innerhalb der Grenzen 0 und $dz = 1$ genommen. Daher ist B gleich der Differenz der Ordinaten $y_1 - y_0 = y_1 - a$, und C ist gleich der Differenz der ersten Ableitungen dieser Ordinaten nach z , E und G gleich der Differenz der 3. resp. der 5. Ableitungen derselben Ordinaten, $\dots \dots \dots$ Bezeichnet man diese Differenzen mit $\alpha, \beta, \gamma, \delta, \dots \dots \dots$, so wird a oder:

$$y_0 = A - \frac{\alpha}{2} + \frac{\beta}{12} - \frac{\gamma}{720} + \frac{\delta}{30240} \pm \dots \dots \dots$$

Setzt man nun eine Basis $z_0 z_n$ in n aequidistante Theile zerlegt voraus, von denen jeder Theil gleich $dz = 1$ sei, bezeichne S die Summe der aequidistanten Ordinaten $y_0 + y_1 + y_2 + \dots + y_{n-2} + y_{n-1}$, sei ferner nach gegebener Definition $\alpha = y_n - y_0$,

$$\beta = \frac{dy_n}{dz} - \frac{dy_0}{dz}, \gamma = \frac{d^3y_n}{dz^3} - \frac{d^3y_0}{dz^3}, \dots, \text{so ist}$$

$$S = A - \frac{\alpha}{2} + \frac{\beta}{12} - \frac{\gamma}{720} + \frac{\delta}{30240} \pm \dots \dots \dots$$

Dies ist die Summationsformel von Maclaurin für den Fall eines Incrementes gleich 1 ; für ein beliebiges Increment h erhält derselbe analog die Formel:

$$S = \frac{A}{h} - \frac{\alpha}{2} + \frac{\beta h}{12} - \frac{\gamma h^3}{720} + \frac{\delta h^5}{30240} \mp \dots \dots \dots$$

Erinnert man sich, dass A die Fläche der Curve von z_0 bis z_n ist und denkt man an die Bedeutung von $\alpha, \beta, \gamma \dots$, so ist leicht die Identität der letztern Formel mit der folgenden, nämlich mit der Euler'schen (für $h = 1$)

$$\sum_{z=0}^{z=n-1} y_z = \int_0^n y dz = \frac{1}{2} \left[y_z \right]_0^n + \left[\frac{B(1)}{2!} \frac{dy}{dz} \right]_0^n - \left[\frac{B(2)}{4!} \frac{d^3y}{dz^3} \right]_0^n \pm \dots$$

worin $B(1), B(2), B(3) \dots$ die Bernoulli'schen Zahlen bedeuten, festzustellen.

21. Euler gibt die Formel auf rein analytischem Wege in den Inst. Calcul. Different. p. II c. V: «Investigatio summae serierum ex Termino generali». Sei

$$y = f(x), \text{ dann wird:}$$

$$v = f(x-1) = y - \frac{dy}{dx} + \frac{d^2y}{2! dx^2} - \frac{d^3y}{3! dx^3} \pm \dots$$

Nun ist, wenn man mit A den Werth für $x = 0$ bezeichnet, $\sum v = \sum y - y + A$, und substituirt man diesen Werth in die Gleichung:

$$\sum v = \sum y - \sum \frac{dy}{dx} + \frac{1}{2!} \sum \frac{d^2y}{dx^2} - \frac{1}{3!} \sum \frac{d^3y}{dx^3} \pm \dots$$

so kommt:

$$y - A = \sum \frac{dy}{dx} - \frac{1}{2!} \sum \frac{d^2y}{dx^2} + \frac{1}{3!} \sum \frac{d^3y}{dx^3} \pm \dots$$

Setzt man $\frac{dy}{dx} = z$, so ergibt sich durch Substitution:

$$\sum z = \int z dx + \frac{1}{2!} \sum \frac{dz}{dx} - \frac{1}{3!} \sum \frac{d^2z}{dx^2} \pm \dots + \text{Constante.}$$

Es ist aber ebenso:

$$\sum \frac{dz}{dx} = z + \frac{1}{2!} \sum \frac{d^2z}{dx^2} - \frac{1}{3!} \sum \frac{d^3z}{dx^3} \pm \dots$$

$$\sum \frac{d^2z}{dx^2} = \frac{dz}{dx} + \frac{1}{2!} \sum \frac{d^3z}{dx^3} - \frac{1}{3!} \sum \frac{d^4z}{dx^4} \pm \dots$$

...

Diese Werthe in die Gleichung für $\sum z$ eingesetzt, ergibt die neue Formel:

$$\sum z = \int z dz + \alpha z + \beta \frac{dz}{dx} + \gamma \frac{d^2z}{dx^2} + \dots$$

und zur Bestimmung der Coeffizienten $\alpha, \beta, \gamma, \dots$ ergeben sich die Gleichungen:

$$\left. \begin{array}{l} \alpha - \frac{1}{2} = 0 \\ \beta - \frac{\alpha}{2} + \frac{1}{6} = 0 \\ \gamma - \frac{\beta}{2} + \frac{\alpha}{6} - \frac{1}{24} = 0 \\ \delta - \frac{\gamma}{2} + \frac{\beta}{6} - \frac{\alpha}{24} + \frac{1}{120} = 0. \\ \dots \dots \dots \end{array} \right\} \text{also: } \begin{array}{l} \alpha = \frac{1}{2} \\ \beta = \frac{1}{12} \\ \gamma = 0 \\ \delta = \frac{1}{720} \\ \dots \dots \dots \end{array}$$

Das Fortschreitungsgesetz der Coeffizienten findet Euler nach einer längeren Untersuchung über die Bernoullischen Zahlen, die hier nicht ausgeführt werden soll, als folgendes: $\alpha = \frac{1}{2}$, $\beta = \frac{B(1)}{2!}$, $\gamma = 0$, $\delta = -\frac{B(2)}{4!}$, $\varepsilon = 0, \dots$ und demnach wird seine Summenformel:

$$\begin{aligned} \Sigma z = & \int z \, dx + \frac{1}{2} z + \frac{B(1)}{2!} \cdot \frac{dz}{dx} - \frac{B(2)}{4!} \cdot \frac{d^3 z}{dx^3} + \frac{B(3)}{6!} \cdot \frac{d^5 z}{dx^5} \\ & - \frac{B(4)}{8!} \cdot \frac{d^7 z}{dx^7} \pm \dots + \text{Const.} \end{aligned}$$

Aus dieser von Euler gegebenen Form erhält man sofort durch Subtraktion von z und durch Annahme von Grenzen, wenn man $z = \varphi(x)$ setzt, die folgende:

$$\begin{aligned} \sum_{x=0}^{x=x-1} \varphi(x) = & \int_0^x \varphi(x) \, dx - \frac{1}{2} \left[\varphi(x) \right]_0^x + \left[\frac{B(1) \varphi'(x)}{2!} \right]_0^x - \left[\frac{B(2) \varphi''(x)}{4!} \right]_0^x \\ & + \left[\frac{B(3) \varphi^{(V)}(x)}{6!} \right]_0^x \mp \dots \end{aligned}$$

22. Unter den zahlreichen Anwendungen, die Euler von dieser Formel macht, findet sich (im nämlichen Kapitel, Art. 157) auch diejenige zur *Ermittlung eines Näherungswertes für Log I(x+1)**. Ist $z = \text{Log } x$, so wird:

$$\sum_{x=1}^{x=x} \text{Log } x = x \text{ Log } x - x + \frac{1}{2} \text{ Log } x + \frac{B(1)}{1 \cdot 2 \cdot x} - \frac{B(2)}{3 \cdot 4 x^3} \pm \dots + C.$$

und für $x = 1$, folgt

$$C = 1 - \frac{B(1)}{1 \cdot 2} + \frac{B(2)}{3 \cdot 4} - \frac{B(3)}{5 \cdot 6} \pm \dots$$

*) Die folgende Darstellung gibt übrigens schon Maclaurin mittelst seiner Summationsformel, v. Treatise of fluxions, art. 842.

Nun ist nach der Formel von Wallis:

$$\frac{\pi}{2} = \frac{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \dots \dots \dots (2x-2)2x}{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \dots \dots \dots (2x-1)(2x-1)} \quad (\text{für } x = \infty)$$

somit

$$\begin{aligned} \log \pi - \log 2 &= 2 \log 2 + 2 \log 4 + 2 \log 6 + \dots + \log 2x \\ &\quad - 2 \log 1 - 2 \log 3 - 2 \log 5 \dots \dots \end{aligned}$$

Weil aber für $\lim x = \infty$:

$$\sum_{x=1}^{x=x} \log x = C + \left(x + \frac{1}{2}\right) \log x - x$$

$$\sum_{x=1}^{x=2x} \log x = C + \left(2x + \frac{1}{2}\right) \log 2x - 2x$$

$$\sum_{x=1}^{x=x} \log 2x = C + \left(x + \frac{1}{2}\right) \log x + x \log 2 - x,$$

so folgt aus den beiden letzten Gleichungen:

$$\log 1 + \log 3 + \log 5 + \dots \log(2x-1) = x \log x + \left(x + \frac{1}{2}\right) \log 2 - x,$$

also für $\lim x = \infty$:

$$\begin{aligned} \log \frac{\pi}{2} &= 2C + (2x + 1) \log x + 2x \log 2 - \log 2 - \log x - 2x \\ &\quad - 2x \log x - (2x + 1) \log 2 + 2x \\ \log \frac{\pi}{2} &= 2C - 2 \log 2, \quad C = \frac{1}{2} \log 2 \pi. \end{aligned}$$

Es ergibt sich somit für

$$\lim_{x=\infty} \sum_{x=1}^{x=x} \log x = \frac{1}{2} \log 2\pi + \left(x + \frac{1}{2}\right) \log x - x, \quad \text{oder}$$

$$\lim_{x=\infty} x! = \sqrt{2\pi} + x^{x+\frac{1}{2}} e^{-x}.$$

23. Die Summationsformel von Euler und Maclaurin ist aber nicht nur geeignet für die Darstellung eines Näherungswertes für $\log \Gamma(x+1)$, sondern auch *zweckmäßig zur Summation der binomischen Terme* in derjenigen Form, in der sie nach Anwendung der sog. Stirling'schen Formel bei der Darstellung des Bernoulli'schen Theorems erscheinen, und in der That ist seit Laplace, der jene Formel von Euler und Maclaurin zuerst für den bezeichneten Zweck verwendete*), kein anderes Summationsverfahren gefunden worden. *Jene Formel ersetzt somit in hinreichender Weise die mühsamen empirischen Methoden Moivre's zur Ermittlung eines Näherungswertes für den Bernoulli-schen Summenausdruck.*

*) S. Note 1 im Anhang.

Der geniale Laplace hat zum ersten Male mittelst seiner «fonctions génératrices» eine noch allgemeinere Methode angegeben, um einen Näherungswert für $\log I(x+1)$ zu erhalten, nach welcher auch die Constante ohne Benutzung der Wallisischen Formel direct aus der Entwicklung hervorgeht*); er hat auch, nach dem Vorgange von Lagrange, die Euler-Maclaurin'sche Summationsformel auf anderem Wege gefunden. Aber Laplace räumt seinen «fonctions génératrices» gewiss einen zu grossen Einfluss auf die Darstellung des Bernoulli'schen Theorems ein, wenn er schreibt**): «Le calcul des fonctions génératrices, appliqué à cet objet, non seulement démontre avec facilité ce théorème, mais de plus il donne la probabilité que le rapport des évènemens observés ne s'écarte que dans certaines limites du vrai rapport de leurs possibilités respectives»; denn alle diese Consequenzen sind in genügend allgemeiner Weise schon mit Hülfe der Formel von Euler und Maclaurin zu ziehen. Schon vor Laplace, um die Mitte des vorigen Jahrhunderts, wäre es möglich gewesen, dem Bernoulli'schen Theorem diejenige analytische Form zu geben, die es heute besitzt. Der Grund, warum es nicht geschehen, liegt darin, dass sich von Moivre bis auf Laplace kein Mathematiker in productiver Weise auf diesem Gebiete bethätigte.

* * *

24. Die Ergebnisse des historischen Theiles dieser Arbeit, der die Entwicklungsgeschichte des Bernoulli'schen Summenausdruckes zum Laplace'schen Integralausdruck geben sollte, fassen wir folgendermassen zusammen:

1. Jakob Bernoulli I. hat nicht versucht, einen Näherungswert für

$$\begin{aligned} m &= \mu p + 1 \\ &\sum \frac{\mu!}{m! n!} p^m q^n \\ m &= \mu p - 1 \end{aligned}$$

zu geben. Weil er das nach ihm benannte Theorem nur als Helfersatz seiner Theorie der Wahrscheinlichkeit *a posteriori* betrachtete, genügte ihm der ganz allgemein gegebene Nachweis, dass mit der Vermehrung der Beobachtungen auch die Wahrscheinlichkeit immer grösser wird, dass die Erfahrungswahrscheinlichkeit eines Ereignisses gleich seiner absoluten wird.

*) Vgl. Note 4 im Anhang.

**) *Essai philosophique sur les probabilités* p. 74. *Théorie anal. des probab.*, introduction p. XLVIII.

2. Abraham de Moivre gab im Prinzip die Laplace'sche Analyse des Bernoulli'schen Theorems. Er fand nicht nur Näherungswerte für den Binomialcoefficienten und für $\Gamma(x)$, sondern gab auch das Laplace'sche Integral als Summe des Bernoulli'schen Ausdrucks in der Form von

$$\frac{2(p+q)}{\sqrt{2pq\mu\pi}} \int_0^1 e^{\frac{p+q}{2pq\mu}x^2} dx.$$

3. James Stirling hat, auf Anregung Moivre's, den cyklometrischen Charakter der den Näherungswert für $\Gamma(x)$ und das Laplace'sche Integral begleitenden Constanten erkannt.

4. Aber erst der Summationsformel, welche von Maclaurin, dann von Euler gefunden worden ist, verdankt das Bernoulli'sche Theorem die allgemeine Entwicklung jener exakten analytischen Form, die ihm von Laplace gegeben wurde.

VII.

25. Der jetzt folgende Abschnitt gibt eine Verallgemeinerung der Serret'schen Ableitung der Stirling'schen Formel.

Die ersten Darsteller dieser Formel benutzten zur Bestimmung der Constanten die Formel von Wallis. Nun hat J. A. Serret in einem Mémoire sur l'évaluation approchée du produit $1 \cdot 2 \cdot 3 \dots x$, lorsque x est un très grand nombre, et sur la formule de Stirling*) auf elegante Weise gezeigt, dass die Formel von Wallis zur Ableitung derjenigen von Stirling vollkommen hinreichend ist. Er sagt darüber einleitend: « Or, cette simple formule de Wallis suffit, à elle seule, pour établir complètement celle de Stirling et la déduction est si facile que la deuxième formule peut être regardée avec raison comme une transformée de la première.» Serret's Darstellung ist die folgende:

Die Formel von Wallis ist:

$$\frac{\pi}{2} = \frac{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \dots (2x-2)(2x-2) 2x}{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \dots 2(x-3)(2x-1)(2x-1)}, \text{ (für } x=\infty\text{)}$$

und sie nimmt die sehr einfache Form**) an:

*) Comptes rendus hebdomadaires des séances de l'Académie des sciences, année 1860, t. I. p. 1662.

**) Die Transformation ergibt zunächst:

$$S(x) = \frac{2}{\pi} \frac{[(x-1)!]^4 2^{4(x-1)}}{[(2x-1)!]^2} \cdot \frac{2x}{[(2x)!]^2},$$

dann nach einfacher Umformung

$$S(x) = \left[\frac{x!}{x^x \sqrt{2\pi x}} \right]^4 : \left[\frac{(2x)!}{(2x)^{2x} \sqrt{4\pi x}} \right]^2 = \frac{[\varphi(x)]^2}{\varphi(2x)}.$$

1. $\frac{[\varphi(x)]^2}{\varphi(2x)} = 1 \quad (\text{für } x = \infty)$

wenn man mit $\varphi(x)$ entweder den Ausdruck:

$$\frac{1 \cdot 2 \cdot 3 \cdot \dots \cdot x}{\sqrt{2\pi} x^{x+\frac{1}{2}}}$$

oder das Produkt dieses Quotienten mit einer Exponentialfunktion von der Form a^x bezeichnet, wobei a eine beliebige positive Constante bedeutet. Die Gleichung 1) gilt also auch, wenn man setzt:

2. $\varphi(x) = \frac{1 \cdot 2 \cdot 3 \cdot \dots \cdot x}{\sqrt{2\pi} e^{-x} x^{x+\frac{1}{2}}} \cdot (e = \text{Basis der natür. Logarith.})$

Aus dieser Gleichung folgt:

3. $\frac{\varphi(x)}{\varphi(x+1)} = \frac{1}{1} \left(1 + \frac{1}{x}\right)^{x+\frac{1}{2}} = e^{-1 + (x+\frac{1}{2}) \text{Log}(1+\frac{1}{x})}.$

Da $x > 1$, wird, wenn Θ' und Θ'' zwei Größen bezeichnen, die sich zwischen 0 und 1 bewegen,

$$\text{Log} \left(1 + \frac{1}{x}\right) = \frac{1}{x} - \frac{\Theta'}{2x^2} = \frac{1}{x} - \frac{1}{2x^2} + \frac{\Theta''}{3x^3}$$

folglich

$$\left(x + \frac{1}{2}\right) \text{Log} \left(1 + \frac{1}{x}\right) = 1 + \left(\frac{\Theta''}{3} - \frac{\Theta'}{4}\right) \frac{1}{x^2} = 1 + \frac{\Theta}{x^2},$$

wo Θ zwischen -1 und $+1$ gelegen ist, daher

$$\frac{\varphi(x)}{\varphi(x+1)} = e^{\frac{\Theta}{x^2}}.$$

Aendert man nun successive x in $x+1, x+2, \dots, 2x-1$, und bezeichnet man mit $\Theta_1, \Theta_2, \Theta_3, \dots, \Theta_{x-1}$ Größen, die zwischen -1 und $+1$ liegen, so wird

$$\frac{\varphi(x+1)}{\varphi(x+2)} = e^{\frac{\Theta_1}{(x+1)^2}}, \frac{\varphi(x+2)}{\varphi(x+3)} = e^{\frac{\Theta_2}{(x+2)^2}}, \dots, \frac{\varphi(2x-1)}{\varphi(2x)} = e^{\frac{\Theta_{x-1}}{(2x-1)^2}}.$$

Multipliziert man alle diese Gleichungen und beobachtet, dass

$$\frac{\Theta_0}{x^2} + \frac{\Theta_1}{(x+1)^2} + \dots + \frac{\Theta_{x-1}}{(2x-1)^2} < \frac{1}{x},$$

so kann man schreiben:

$$\frac{\varphi(x)}{\varphi(2x)} = e^{\frac{\Theta}{x}},$$

wo Θ eine Größe ist, die zwischen -1 und $+1$ liegt, und wird $x = \infty$, so hat man

4. $\frac{\varphi(x)}{\varphi(2x)} = 1. \quad (\text{für } x = \infty).$

Dividirt man nun Gleichung 1) durch 4) so kommt:

$$\varphi(x) = 1 \quad (\text{für } x = \infty)$$

d. h. nach Formel 2):

$$1 \cdot 2 \cdot 3 \dots x = \sqrt{2\pi} e^{-x} x^{x+\frac{1}{2}} (1 + \varepsilon_x),$$

wo ε_x eine Grösse ist, die für $x = \infty$ zu 0 wird.

26. Ist nun diese von Serret gegebene Darstellung eines Näherungswertes für $\Gamma(x+1)$ auch die einfachste und eleganteste, die je gegeben wurde, so erscheint sie doch einer Verallgemeinerung fähig zu sein. Wenn man die von Serret gefundene Funktion mit $S(x)$ bezeichnet, so ergibt sich aus der Formel von Wallis für

$$\lim_{x \rightarrow \infty} S(x) = \left[\frac{x!}{x^x \sqrt{2\pi x}} \right]^4 : \left[\frac{(2x)!}{(2x)^{2x} \sqrt{4\pi x}} \right]^2 = 1,$$

oder wenn man den Ausdruck

$$\frac{x!}{x^x \sqrt{2\pi x}}$$

mit $\varphi(x)$ bezeichnet, so wird

$$1) \lim_{x \rightarrow \infty} S(x) = \frac{\varphi^2(x)}{\varphi(2x)} = 1.$$

Serret setzt aber die Funktion

$$\frac{x!}{x^x \sqrt{2\pi x} e^{-x}} = \varphi(x).$$

Diese Erweiterung von $S(x)$ mit e^{-x} ist in der That beim Gedanken an den Stirling'schen Näherungswert für $x!$ sehr naheliegend.

Aber im allgemeineren Falle *muss jene Exponentialgrösse erst im Verlaufe der Entwicklung als gewisse Bedingung sich darstellen*, wie im Folgenden gezeigt werden soll.

Es sei also

$$\varphi(x) = \frac{x!}{x^x \sqrt{2\pi x}},$$

$$\lim_{x \rightarrow \infty} S(x) = \frac{\varphi^2(x)}{\varphi(2x)} = 1.$$

Wie finden wir hieraus einen Werth für $x!$? Offenbar, wenn es gelingt, nachzuweisen, unter welcher Bedingung

$$\lim_{x \rightarrow \infty} \frac{\varphi(x)}{\varphi(2x)} = 1 \text{ ist. Denn alsdann wird}$$

$$\lim_{x \rightarrow \infty} \frac{\frac{\varphi^2(x)}{\varphi(2x)}}{\frac{\varphi(x)}{\varphi(2x)}} \Rightarrow \varphi(x) = \frac{x!}{x^x \sqrt{2\pi x}} = 1.$$

Es soll nun untersucht werden, unter welchen Bedingungen

$$\lim_{x \rightarrow \infty} \frac{\varphi(x)}{\varphi(2x)} = 1 \text{ ist.}$$

Dazu ist

$$\frac{\varphi(x)}{\varphi(x+1)} = \left(1 + \frac{1}{x}\right)^{x+\frac{1}{2}}$$

$$\text{und } \log \frac{\varphi(x)}{\varphi(x+1)} = 1 + \frac{1}{12x^2} - \frac{1}{12x^3} + \dots + \frac{(n-1)(-1)^n}{2n(n+1)x^n} + \dots \text{ in inf.}$$

Die Reihe ist, da $x > 1$, convergent und die Glieder nehmen, selbst für $x = 1$, schon vom Gliede $-\frac{1}{12x^3}$ an, ab.

Es wird daher

Werden die Gleichungen 2) addiert, so findet man

$$3) \log \frac{\varphi(x)}{\varphi(2x)} = x + \frac{\theta}{x}, \text{ wo } 0 < \theta < \frac{1}{12}$$

woraus, wenn man zur Exponentialfunktion übergeht, folgt

$$4) \lim_{x \rightarrow \infty} \frac{\varphi(x)}{\varphi(2x)} e^{-x} = e^{\frac{\Theta}{x}}, \text{ oder}$$

Die Multiplikation mit der Exponentialgrösse e^{-x} ist somit die gesuchte Bedingung für die Existenz der Gleichung:

$$\lim_{x \rightarrow \infty} \frac{\varphi(x)}{\varphi(2x)} = 1.$$

Durch Division der Gleichungen 1) und 4) geht hervor:

$$\lim_{x \rightarrow \infty} \varphi(x) e^x = 1, \text{ oder da } \varphi(x) = \frac{x!}{x^x \sqrt{2\pi x}},$$

resultiert schliesslich

$$\lim_{x \rightarrow \infty} x! = x^x e^{-x} \sqrt{2\pi x},$$

oder $x! = x^x e^{-x} \sqrt{2\pi x} (1+\omega_x)$,

worin ω_x für $\lim x = \infty$ verschwindet.

Bei der Bestimmung der Grenzen findet man*)

$$x^x e^{-x} \sqrt{2\pi x} < x! < x^x e^{-x} \sqrt{2\pi x} + \frac{1}{12} \sqrt{2\pi x} \quad \text{und}$$

weil die Quantität $\frac{1}{12x}$ schon in der Gleichung 3) als Grenze aufgetreten, hegte ich die Vermuthung, dass sie sich ebenfalls durch die obige Entwicklung als Grenze finden liesse. Der Nachweis ist mir aber bis jetzt nicht gelungen.

VIII.

27. Dieser Abschnitt gibt einen neuen vereinfachten Ausdruck für das Bernoullische Theorem.

Es wurde im historischen Theil dieser Arbeit gezeigt, wie Moivre zuerst für den Bernoullischen Summenausdruck

$$W = \sum_{m=\mu p-1}^{m=\mu p+1} \frac{\mu!}{m! n!} p^m q^n,$$

worin $m + n = \mu$, $p + q = 1$ und $1 = \gamma \sqrt{2\pi p q m}$ ist, einen Integralausdruck gegeben hat, welchen alsdann Laplace wie in Note 1 im Anhang ersichtlich, mit vollkommeneren Methoden genauer gab durch

$$W = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-t^2} dt + \frac{e^{-\gamma^2}}{\sqrt{2\pi \mu p q}}.$$

Dieser Ausdruck ist seit Laplace unverändert geblieben, man findet ihn heute noch in den besten Handbüchern für Wahrscheinlichkeitsrechnung, so in denen von Meyer und Czuber, von Bertrand u. a. m.

Bei Operationen mit demselben erweist sich jedoch die Restfunktion $\frac{e^{-\gamma^2}}{\sqrt{2\pi \mu p q}}$ als sehr unbequem. Um so mehr muss es auffallen, dass seit Laplace noch niemand es versucht hat, dieselbe durch Vereinigung mit dem Integral ihrer isolirten Stellung zu entheben.

Dass dies möglich ist, soll im Folgenden gezeigt werden.

Es sei

$$y = \frac{\mu}{m! n!} p^m q^n \text{ das allgemeine Glied des}$$

Binoms $(p + q)^\mu$, worin p und q die bekannte Bedeutung haben. Als dann wird, wie es schon Laplace gezeigt hat, mit Hülfe der Formel

*) Serret gibt diese Grenzenbestimmung auf hübsche Weise in seinen Cours d'algèbre supérieure (5. éd., Paris 1885) tome II, art. 393, p. 218.

von Stirling und unter Berücksichtigung des Satzes, dass diejenige Combination der Zahlen des Eintreffens und Nichteintreffens des Ereignisses ein Maximum von Wahrscheinlichkeit besitzt, die unter der Relation steht $p : q = m : n$, die Wahrscheinlichkeit, dass bei μ Versuchen das Ereigniss (dessen einfache und konstante Wahrscheinlichkeit gleich p , dessen entgegengesetzte gleich q ist) eine Anzahl Male eintreffe, die zwischen $\mu p \pm 1$ liegt, ausgedrückt durch

$$W = \sum_{\substack{m=\mu p+1 \\ m=\mu p-1}} \frac{\mu!}{m! n!} p^m q^n = y_{\nu+1} + y_{\nu+1+1} + \dots$$

$$+ y_{\nu+1} + y_{\nu} + y_{\nu+1} + \dots + y_{\nu+1-1} + y_{\nu+1},$$

worin also in allen Gliedern m durch μp und n durch μq ersetzt ist und y_{ν} das Maximalglied bedeutet.

Dann kann man setzen :

$$W = \sum_{\substack{\lambda=0 \\ \lambda=1}}^{1} [y_{\nu-\lambda} + y_{\nu+\lambda}] = y_{\nu}, \text{ oder}$$

$$= \sum_{\lambda=0}^1 \varphi(\lambda) - \frac{1}{2} \varphi(0), \text{ wenn}$$

$$\varphi(\lambda) = \frac{2}{\sqrt{2\pi\mu pq}} \cdot e^{-\frac{\lambda^2}{2\mu pq}}, \text{ also } \varphi(0) = \frac{2}{\sqrt{2\pi\mu pq}} \text{ ist.}$$

28. So viel mir bekannt ist, wurde der Uebergang von der zuletzt gegebenen Summe zum Integral seit Laplace immer mit Hilfe der Summationsformel von Maclaurin und Euler gemacht. Eine eigene Methode für diesen Uebergang, auf mechanischer Quadratur beruhend, gab mein verehrter Lehrer, Herr Privatdozent *Dr. Ch. Moser* in Bern, der sich bei versicherungstechnischen Arbeiten oft mit dieser Materie beschäftigte, in seiner Vorlesung über das Bernoullische Theorem (im Sommer-S. 1892) und zwar in folgender Weise :

$$\text{Sei } f(x) = \frac{2}{\sqrt{2\pi\mu pq}} e^{-\frac{x^2}{2\mu pq}} \text{ und}$$

$$W = \sum_{x=0, 1, 2, \dots} f(x) - \frac{1}{2} f(0)$$

Die Funktion $f(x)$ liefert, weil μpq positiv ist, für $x = 0$ ein Maximum und nimmt mit wachsendem x stetig ab. Die rechte Seite

der Gleichung für W kann, da x nur die ganzzahligen Werthe 0, 1, 2, 1 betritt, geschrieben werden:

$$W = \frac{1}{2}f(0) + 1 \cdot f(1) + 1 \cdot f(2) + \dots + 1 \cdot f(l).$$

Die einzelnen Summanden seien als Rechtecke dargestellt und zwar $f(0)$ mit der Basis $\frac{1}{2}$ und die übrigen Werthe je mit der Basis 1.

Werden diese Rechtecke über einer gemeinsamen Grundlinie aneinander gereiht und wird über dieser Grundlinie als Axe der x die Curve $f(x)$ construirt, so schneidet diese die der Basis gegenüberliegenden Seiten der für $f(1), f(2) \dots f(l)$ erstellten Rechtecke je in der Mitte. Die Fläche, welche von der Grundlinie, den Ordinaten $f(0)$ und $f\left(1 + \frac{1}{2}\right)$ und der Curve $f(x)$ eingeschlossen ist,

hat zum Ausdrucke: $\int_0^{1+\frac{1}{2}} f(x) dx$. Substituirt man diese Fläche für die Summe der Rechtecke, so kommt bei einem einzelnen Rechteck ungefähr ein so grosses Dreieck hinzu, wie die Curve von dem Rechteck abschneidet, — absolut genau, sobald die Curve für ihren über der Basis eines Rechtecks gelegenen Theil als geradlinig betrachtet werden kann. Nur beim ersten Rechteck, $\frac{1}{2}f(0)$, hebt sich das kleine Fehlerdreieck nicht auf. Dieses wird jedoch, da $f(x)$ für $x = 0$ ein Maximum aufweist, sehr klein. In Näherung muss daher gelten:

$$W = \sum_{x=0, 1, 2, \dots, l} f(x) - \frac{1}{2} f(0) = \int_0^{1+\frac{1}{2}} f(x) dx.$$

Das Resultat, das diese geometrische Ueberlegung liefert, leuchtete mir ein und regte mich an, eine Untersuchung auf analytischem Wege vorzunehmen.

29. Sei also

$$1) W = \sum_{x=0}^{x=1} \varphi(x) - \frac{1}{2} \varphi(0) \text{ oder auch}$$

$$2) W = \sum_{x=0}^{x=1-1} \varphi(x) + \varphi(1) - \frac{1}{2} \varphi(0),$$

worin $\varphi(x) = \frac{2}{\sqrt{2\pi\mu pq}} e^{-\frac{x^2}{2\mu pq}}$ ist.

Nach Euler's Summationsformel*):

$$\sum_{x=0}^x \varphi(x) = \int_0^{x+1} \varphi(x) dx - \frac{1}{2} \left[\varphi(x) \right]_0^{x+1} + \left[\frac{B(1)\varphi'(x)}{2!} \right]_0^{x+1} - \left[\frac{B(2)\varphi''(x)}{4!} \right]_0^{x+1} \pm \dots$$

wird

$$3) \quad \sum_{x=0}^l \varphi(x) = \int_0^{l+1} \varphi(x) dx - \frac{1}{2} \varphi(l) + \frac{1}{2} \varphi(0)$$

und

$$4) \quad \sum_{x=0}^{l-1} \varphi(x) = \int_0^l \varphi(x) dx - \frac{1}{2} \varphi(l) + \frac{1}{2} \varphi(0),$$

bei Vernachlässigung der mit $\varphi'(x)$, $\varphi''(x)$. . . behafteten Glieder, weil $\varphi'(x) = -\frac{4}{\sqrt{2\pi\mu pq}} \cdot \frac{1}{2\mu pq} e^{-\frac{x^2}{2\mu pq}}$, also von der Ordnung $\frac{1}{\mu}$ ist (wo μ sehr gross vorausgesetzt wird).

Die Werthe 3) und 4) in die Gleichungen 1) und 2) substituirt, ergibt :

$$5) \quad W = \int_0^{l+1} \varphi(x) dx - \frac{1}{2} \varphi(l+1) \text{ oder}$$

$$6) \quad W = \int_0^l \varphi(x) dx + \frac{1}{2} \varphi(l).$$

Also liegt

W zwischen $\int_0^l \varphi(x) dx$ und $\int_0^{l+1} \varphi(x) dx$, und es sei daher

*) Dr. Bruno Borchardt benutzt in seiner «Einführung in die Wahrscheinlichkeitslehre» (Berlin 1889, Jul. Springer) diese Summationsformel unrichtig, indem er die Grenzen auf beiden Seiten der Gleichung gleich nimmt, während die obere Grenze rechts um eine Einheit höher genommen werden muss. Borchardt erhält auch ein unrichtiges Resultat (p. 31 und 32 seines Buches). Auch in Meyer und Czuber «Vorlesungen über Wahrscheinlichkeit» (Leipzig 1879, Teubner) finden sich Unrichtigkeiten (oder sind es bloss unkorrigierte Druckfehler?) im Gebrauche der Grenzen (p. 101 und 102).

$W = \int_0^{1+\varepsilon} \varphi(x) dx$, wo ε eine kleine Grösse, zwischen 0 und 1 gelegen, ist, die bestimmt werden soll. Zu diesem Zwecke suche ich zu entwickeln

$$\int_1^{1+\varepsilon} \varphi(x) dx.$$

7) Man setze $\int \varphi(l) = f(l)$, dann wird

8) $\int_1^{1+\varepsilon} \varphi(x) dx = f(l+\varepsilon) - f(l)$ oder nach Taylor

$$= f(l) + \varepsilon f'(l) + \frac{\varepsilon^2 f''(l)}{2!} + \frac{\varepsilon^3 f'''(l)}{3!} + \dots \text{ in inf.} - f(l)$$

$$= \varepsilon f'(l) + \frac{\varepsilon^2 f''(l)}{2!} + \frac{\varepsilon^3 f'''(l)}{3!} + \dots \text{ in inf.}$$

Es ist aber nach 7):

$$f'(l) = \varphi(l), f''(l) = \varphi'(l), f'''(l) = \varphi''(l), \dots, \dots, \dots,$$

somit folgt durch Substitution dieser Werthe in 8):

$$9) \int_1^{1+\varepsilon} \varphi(x) dx = \varepsilon \varphi(l) + \frac{\varepsilon^2 \varphi'(l)}{2!} + \frac{\varepsilon^3 \varphi''(l)}{3!} + \dots \text{ in inf.}$$

Weil $\varepsilon < 1$ ist, so ist die Reihe q) convergent und man erhält unter Vernachlässigung der Glieder von der Ordnung $\frac{1}{\mu}$ in erster Näherung:

$$\int_1^{1+\varepsilon} \varphi(x) dx = \varepsilon \varphi(l).$$

Es war aber nach 6):

$$\int_1^{1+\varepsilon} \varphi(x) dx = \frac{1}{2} \varphi(l), \text{ also}$$

$$\varepsilon \varphi(l) = \frac{1}{2} \varphi(l) \quad \text{oder}$$

$$\varepsilon = \frac{1}{2}.$$

Daher wird:

$$W = \int_0^1 \varphi(x) dx + \int_1^{1+\frac{1}{2}} \varphi(x) dx = \int_0^{1+\frac{1}{2}} \varphi(x) dx.$$

Analog folgt aus Gleichung 5):

$\int_0^{l+1-\theta} \varphi(x) dx = -\frac{1}{2} \varphi(l+1)$, und durch Entwicklung des Integrals links, nach Taylor, wie oben, ergibt sich wieder:

$$-\theta \varphi(l+1) = -\frac{1}{2} \varphi(l+1) \quad \text{oder}$$

$$\theta = \frac{1}{2}$$

und

$$W = \int_0^{l+1} \varphi(x) dx - \int_{l+1-\frac{1}{2}}^{l+1} \varphi(x) dx = \int_0^{l+\frac{1}{2}} \varphi(x) dx.$$

30. Es war aber $\varphi(x) = \frac{2}{\sqrt{2\pi\mu pq}} e^{-\frac{x^2}{2\mu pq}}$, also wird

$$W = \frac{2}{\sqrt{2\pi\mu pq}} \int_0^{l+\frac{1}{2}} e^{-\frac{x^2}{2\mu pq}} dx.$$

Oder setzt man $\xi^2 = \varrho x^2$, wo $\varrho = \frac{1}{2\mu pq}$, so wird

$$\xi = x \sqrt{\varrho}, \quad d\xi = dx \sqrt{\varrho} \quad \text{und} \quad dx = \frac{d\xi}{\sqrt{\varrho}}.$$

Für die Grenzen gilt dann:

$$x = 0 \quad , \quad \xi = 0$$

$$x = l + \frac{1}{2} \quad , \quad \xi = \left(l + \frac{1}{2}\right) \sqrt{\varrho} = \gamma$$

und nach einfacher Substitution geht hervor:

$$W = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-\xi^2} d\xi.$$

Die dem Laplace'schen Integralausdrucke anhaftende, bei Anwendungen desselben lästig werdende Restfunktion ist hier mit dem Integral vereinigt. Dabei hat eine Veränderung der oberen Grenze stattgefunden: im Laplace'schen Integral war $\gamma = l \sqrt{\frac{1}{2\mu pq}}$, hier

$$\text{ist } \gamma = \left(l + \frac{1}{2}\right) \sqrt{\frac{1}{2\mu pq}}.$$

Dieser neue Ausdruck erleichtert nicht nur die sehr zahlreichen theoretischen und praktischen Anwendungen des Bernoulli'schen Theorems, sondern ermöglicht auch genauere Resultate, und ich behalte mir vor, gelegentlich einige dieser Consequenzen zu ziehen.

Anhang.

Note 1. Laplace gibt folgende Darstellung des Bernoulli'schen Theorems*): Seien p und q resp. die einfachen Wahrscheinlichkeiten der Ereignisse E und E' , dann ist die Wahrscheinlichkeit, dass in $m + n = \mu$ Versuchen das Ereigniss E m mal, E' n mal eintreffe, gleich dem $(m + 1)^{\text{ten}}$ Terme in der Entwicklung von $(p + q)^\mu$, nämlich gleich

$$\cdot \frac{1 \cdot 2 \cdot 3 \dots \mu}{1 \cdot 2 \cdot 3 \dots m \cdot 1 \cdot 2 \cdot 3 \dots n} p^m q^n.$$

Bezeichnen wir den grössten Term in dieser Entwicklung mit M , so wird sein ihm vorangehender gleich $\frac{Mp}{q} \cdot \frac{n}{m+1}$, sein nachfolgender gleich $\frac{Mq}{p} \cdot \frac{m}{n+1}$ sein. Damit aber M der grösste Term ist, muss gelten

$$\frac{m}{n+1} < \frac{p}{q} < \frac{m+1}{n}$$

und hieraus folgt, dass

$$(\mu+1) p - 1 < m < (\mu+1) p$$

oder $m = (\mu+1) p - \sigma$, wo $\sigma < 1$, ist.

Nun wird

$$p = \frac{m+\sigma}{\mu+1}, \quad q = 1-p = \frac{n+1-\sigma}{\mu+1}, \quad \frac{p}{q} = \frac{m+\sigma}{n+1-\sigma},$$

und sind m und n sehr grosse Zahlen, so gilt die Relation

$$\frac{p}{q} = \frac{m}{n},$$

d. h. das Eintreffen derjenigen Combination der Ereignisse E und E' hat ein Maximum von Wahrscheinlichkeit, die unter der Relation $p:q = m:n$ steht.

*) Théorie analytique des probabilités (3. éd. Paris 1820) Liv. II, Chap. II, p. 280 e. l. s.

Der 1^{te} Term nach dem grössten M ist gleich

$$\frac{\mu!}{(m-l)!(n+l)!} p^{m-l} q^{n+l}.$$

Nun ist

$$1 \cdot 2 \cdot 3 \dots n = n^{n+\frac{1}{2}} \sqrt{2\pi} \left\{ 1 + \frac{1}{12n} + \dots \right\}$$

und es wird

$$\frac{1}{(m-l)!} = (m-l)^{l-m-\frac{1}{2}} \frac{e^{m-l}}{\sqrt{2\pi}} \left\{ 1 - \frac{1}{12(m-l)} - \dots \right\}$$

$$\frac{1}{(n+l)!} = (n+l)^{l-n-\frac{1}{2}} \frac{e^{n+l}}{\sqrt{2\pi}} \left\{ 1 - \frac{1}{12(n+l)} - \dots \right\}$$

Durch logarithmische Entwicklung und unter Vernachlässigung der Glieder

von der Ordnung $\frac{1}{\mu}$ wird

$$(m-l)^{l-m-\frac{1}{2}} = e^{l-\frac{l^2}{2m}} m^{l-m-\frac{1}{2}} \left\{ 1 + \frac{1}{2m} - \frac{l^3}{6m^2} \right\}$$

$$(n+l)^{l-n-\frac{1}{2}} = e^{l-\frac{l^2}{2n}} n^{l-n-\frac{1}{2}} \left\{ 1 + \frac{1}{2n} - \frac{l^3}{6n^2} \right\}.$$

Weil $p = \frac{m+s}{\mu+1}$ ist ($s < 1$), so kann man setzen: $p = \frac{m-\zeta}{\mu}$, wenn

ζ sich in den Grenzen $\frac{n}{\mu+1}$ und $-\frac{\mu-n}{\mu+1}$ bewegt, also ein ächter Bruch ist. Dann wird $q = \frac{n+\zeta}{\mu}$ und man hat

$$p^{m-l} q^{n+l} = \frac{m^{m-l} n^{n+l}}{\mu^\mu} \left\{ 1 + \frac{\mu \zeta l}{mn} \right\},$$

woraus sich ergibt

$$\frac{\mu!}{(m-l)!(n+l)!} p^{m-l} q^{n+l} = \frac{\sqrt{\mu} e^{-\frac{\mu l^2}{2mn}}}{\sqrt{2\pi mn}} \left\{ 1 + \frac{\mu \zeta l}{mn} + \frac{l(n-m)}{2mn} - \frac{l^3}{6m^2} + \frac{l^3}{6n^2} \right\}.$$

Nimmt man in der letzten Gleichung l negativ, so erhält man einen Ausdruck für den Term, der dem grössten um 1 Glieder vorausgeht, und die Summe der beiden ist gleich

$$\frac{2\sqrt{\mu}}{\sqrt{2\pi mn}} e^{-\frac{\mu l^2}{2mn}}$$

Nun wird die Summe derjenigen Terme in der Entwicklung von $(p+q)^\mu$, welche gelegen sind zwischen 2 Termen, die nach links und rechts aequidistant um 1 Terme vom grössten M abstehen (inclus. die äussersten), ausgedrückt durch das endliche Integral:

$$\sum_{l=0}^{l=1} \frac{2\sqrt{\mu}}{\sqrt{2\pi mn}} e^{-\frac{\mu l^2}{2mn}} = \frac{\sqrt{\mu}}{\sqrt{2\pi mn}},$$

wobei berücksichtigt ist, dass man das grösste Glied, welches man für $l = 0$ bekommt, nur einmal zu zählen hat.

Wenn nun y_1 eine Funktion von l bezeichnet, so gilt die Formel (nach Maclaurin und Euler):

$$\sum y_1 = \int y_1 dl - \frac{1}{2} y_1 + \frac{1}{12} \frac{dy}{dl} + \dots + \text{Const.},$$

welche sich in unserm Falle, wo $y_1 = \frac{2\sqrt{\mu}}{\sqrt{2\pi mn}} e^{-\frac{\mu l^2}{2mn}}$ ist, und die erste Derivirte nach l von der Ordnung $\frac{1}{\mu}$ wird und vernachlässigt werden kann, in erster Näherung reduziert auf:

$$\sum y_1 = \int y_1 dl - \frac{1}{2} y_1 + \text{Const.}$$

Und nimmt man rechts die bestimmten Integrale (deren obere Grenze um eine Einheit höher ist als bei der Summe links) so wird, wenn man das Maximalglied für $l = 0$ mit Y bezeichnet:

$$\begin{aligned} \sum_{\lambda=0}^{\lambda=l-1} y_\lambda &= \int_0^l y d\lambda - \frac{1}{2} y_1 + \frac{1}{2} Y \text{ oder auch} \\ \sum_{\lambda=0}^{\lambda=l} y_\lambda &= \int_0^l y d\lambda + \frac{1}{2} y_1 + \frac{1}{2} Y. \end{aligned}$$

Substituirt man nun für y_1 und für Y die gegebenen Werthe in den Ausdruck 1), so wird derselbe, wenn man $t = \frac{1}{\sqrt{2mn}} \sqrt{\frac{\mu}{\mu}}$ setzt, gleich

$$\frac{2}{\sqrt{\pi}} \int_0^{l\sqrt{\frac{\mu}{2mn}}} e^{-t^2} dt + \frac{e^{-t^2} \sqrt{\mu}}{\sqrt{2\pi mn}}$$

Weil nun $m = \mu p + \zeta$, ($\zeta < 1$), so hat man

$$\frac{m+1}{\mu} - p = \frac{1+\zeta}{\mu} = \frac{t\sqrt{2mn}}{\mu\sqrt{\mu}} + \frac{\zeta}{\mu},$$

also drückt die Formel 2) die Wahrscheinlichkeit aus dafür, dass die Differenz zwischen dem Verhältniss der Zahl des Eintreffens des Ereignisses E zu μ , der Gesamtzahl aller Versuche und der einfachen Wahrscheinlichkeit dieses Ereignisses E innerhalb der Grenzen

$$\pm \frac{t\sqrt{2mn}}{\mu\sqrt{\mu}} + \frac{\zeta}{\mu}$$

gelegen ist.

Note 2. In Propos. XXI interpolirt Stirling die Fakultätenreihe 1, 1, 1.2, 1.2.3, 1.2.3.4, und zwar speciell das zwischen 1 und 1 liegende Glied.

Wegen der stark vorhandenen Divergenz der Differenzen der Reihe interpolirt er deren Logarithmenreihe, sucht zunächst den Logarithmenterm zwischen 10! und 11! und findet*) dafür 7.0755259569 dem als Numerus 11899423.08 entspricht. In Propos. XVI hat Stirling aber zugleich gezeigt, dass, wenn die intermediären Glieder der obigen Fakultätenreihe mit a, b, c, d, bezeichnet werden, die Relationen bestehen:

$b = \frac{3}{2}a, c = \frac{5}{2}b, d = \frac{7}{2}c \dots \dots \dots$ Indem er nun das Glied zwischen 10! und 11! successive durch $\frac{19}{2}, \frac{19}{2}, \frac{17}{1}, \dots \dots \frac{3}{2}$ dividirt, erhält er für das gesuchte intermediäre Glied die Zahl 0.8862269251. Das Quadrat dieses Werthes ist gleich der Fläche des Kreises vom Durchmesser 1, also wird das Glied selber gleich $\frac{1}{2}\sqrt{\pi}$ sein. Ebenso folgt hieraus, dass dasjenige intermediäre Glied, das dem ersten vorausgeht, gleich $\sqrt{\pi}$ sein wird.

Stirling findet also durch äusserst mühsame numerische Berechnung folgende Resultate:

$$\Gamma\left(\frac{23}{2}\right) = \frac{21 \cdot 19 \cdot 17 \dots \dots 5 \cdot 3}{2 \cdot 2 \cdot 2 \dots \dots 2 \cdot 2} \Gamma\left(\frac{3}{2}\right) = 11899423.08$$

$$\Gamma\left(\frac{3}{2}\right) = 0.8862269251 = \frac{1}{2}\sqrt{\pi}$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Dieses letzte Resultat benutzt Stirling bei der ersten Lösungsmethode des Coeffizientenproblems in Propos. XXII, die im wesentlichen darin be-

*) Mit Hülfe der Interpolationsformel ($T = \text{allgemeines Glied}$):

$$T = \frac{A + az}{2} + \frac{3B + bz}{2} \frac{z^2 - 1}{4 \cdot 6} + \frac{5C + cz}{2} \frac{(z^2 - 1)(z^2 - 9)}{4 \cdot 6 \cdot 8 \cdot 10} + \dots \dots$$

Die Formel gilt allgemein (auch für die intermediären Glieder) einer Reihe mit 2 Mittelgliedern, von der Form

..... $_3A \ _1A \ A_1 \ A_3 \dots \dots \dots$
wenn die 1. Differenzen $_2a \ a \ a_4 \dots \dots \dots$

die 2. " $_1B \ B_1 \dots \dots \dots$

die 3. " $b \dots \dots \dots$

....., wenn man ferner $A = _1A + A_1, B = _1B + B_1, C = _1C + C_1 \dots \dots$ setzt und mit z das Verhältniss bezeichnet, welches die Entfernung des zu interpolirenden Gliedes T von der Mitte zum constanten Intervall der Variablen hat. Stirling gibt diese Formel in Propos. XX. deutet aber nur an, er sei mit Hülfe der Differenzenrechnung auf dieselbe gekommen.

steht, mit Hülfe der unten gegebenen Interpolationsformel das m^{te} Glied der Reihe

$$1, \frac{2}{1} A, \frac{4}{3} B, \frac{6}{5} C, \frac{8}{7} D \dots \dots$$

zu bestimmen.

Note 3. Die *Inflexionspunkte der Wahrscheinlichkeitscurve* bestimmt Moivre*) wie folgt: Wenn alle Glieder einer binomischen Entwicklung $(a + b)^n$ in gleichen Abständen auf eine gemeinsame Basis aufgetragen werden und man durch die Endpunkte derselben eine Curve legt, so hat diese 2 Inflexionspunkte, die auf verschiedenen Seiten des Maximalgliedes gelegen sind. Um nun den Inflexionspunkt zu bestimmen, sei H die zugehörige Ordinate, deren Stelle vom Anfang der Reihe aus mit l bezeichnet werde, dann wird das nächste Glied gegen den Anfang der Reihe hin gleich

$$\frac{l-1}{n-l+2} \cdot H \cdot \frac{a}{b},$$

und das nächste gegen das Ende der Reihe gleich

$$\frac{n-l+1}{l} \cdot H \cdot \frac{b}{a}.$$

Werden nun die Differenzen dieser Glieder in Bezug auf H gleichgesetzt, so ergibt sich aus

$$\frac{n-l+1}{l} \cdot \frac{b}{a} - 1 = 1 - \frac{l-1}{n-l+2} \cdot \frac{a}{b}$$

als Werth für l

$$l = \frac{a + 3b + 2bn \pm \sqrt{a^2 + 6ab + 4nab + b^2}}{2a + 2b}.$$

Wird im letzten Ausdruck die Wurzel mit r bezeichnet, so wird das Intervall, um welches der Inflexionspunkt links resp. rechts vom grössten Gliede absteht, gleich $\frac{a-b+r}{2a+2b}$ resp. $\frac{b-a+r}{2a+2b}$ sein, und wenn $a = b$ (wenn also die Wahrscheinlichkeitscurve symmetrisch zum grössten Terme verläuft), ist jeder der beiden Inflexionspunkte vom grössten und mittleren Gliede um das Intervall $\frac{1}{2} \sqrt{n+2}$ oder $\frac{1}{2} \sqrt{n}$ (für $n = \text{sehr gross}$) abstehend.

Note 4. Laplace findet auf folgende Weise einen Näherungswert für die Fakultät**): Sei

*) *Miscell. analytica lib. V, c. IV.*

**) *V. Mémoires de l'Académie royale des sciences pour l'année 1778: Mémoires sur les probabilités par P. S. Laplace art. XXIII.* Dort gibt Laplace mittelst des Euler'schen Integrals $\int_0^1 x^p (1-x)^q dx$ auch einen Näherungswert für den Binomialcoeffizienten.

$$y = x^p e^{-x}, \text{ so wird} \\ \int_0^\infty x^p e^{-x} dx = p!$$

y liefert sein Maximum, wenn $x = p$ ist. Setzt man nun $p = \frac{1}{\alpha}$ und $x = \frac{1}{\alpha} + \theta$, so wird

$$\log y - \log p^p e^{-p} = \frac{1}{\alpha} \log (1 + \alpha \theta) - \theta \text{ und} \\ \int_0^\infty y dx = p^p e^{-p} \int e^{\frac{1}{\alpha} \log (1 + \alpha \theta)} - \theta d\theta.$$

Substituieren wir noch

$$\log (1 + \alpha \theta) - \alpha \theta = -\alpha t^2, \text{ so wird} \\ \frac{\alpha \theta^2}{2} - \frac{\alpha^3 \theta^3}{3} + \frac{\alpha^5 \theta^5}{4} - \dots = t^2.$$

Nun kann man finden:

$$\theta = \frac{1}{\sqrt{\alpha}} (ht + h' \alpha^{\frac{1}{2}} t^2 + h'' \alpha t^3 + \dots), \\ \text{worin } h = \sqrt{2}, h' = \frac{2}{3}, h'' = \frac{\sqrt{2}}{18}, \dots$$

und

$$d\theta = \frac{dt}{\sqrt{\alpha}} (h + 2h' \alpha^{\frac{1}{2}} t + 3h'' \alpha t^2 + \dots)$$

Dann wird

$$\int_0^\infty y dx = p^{p+\frac{1}{2}} e^{-p} \int_{-\infty}^\infty (h + 2h' \alpha^{\frac{1}{2}} t + 3h'' \alpha t^2 + \dots) e^{-t^2} dt.$$

Nun ist

$$\int_0^\infty t^{2n} e^{-t^2} dt = \frac{1 \cdot 3 \cdot 5 \dots (2n-1)}{2^n} \int_0^\infty e^{-t^2} dt.$$

und mit Hilfe von

$$\int_0^\infty \int_0^\infty e^{\xi(1+\eta)} d\xi d\eta = \frac{\pi}{2} \text{ findet man, dass}$$

$$\int_0^\infty e^{-t^2} dt = \frac{1}{2} \sqrt{\pi}. \text{ Somit ergibt sich}$$

$$\int_0^\infty t^{2n} e^{-t^2} dt = \frac{1 \cdot 3 \cdot 5 \dots (2n-1)}{2^n} \cdot \frac{1}{2} \sqrt{\pi},$$

und wenn darnach die letzte Formel für $\int_0^\infty y dx$ integriert wird, erhält man schliesslich*):

$$p! = \int_0^\infty y dx = p^{p+\frac{1}{2}} e^{-p} \sqrt{\pi} \left(h + 1 \cdot 3 \frac{a^2 h''}{2} + 1 \cdot 3 \cdot 5 \frac{a^4 h''''}{2^2} + \dots \right)$$

oder

$$p! = p^{p+\frac{1}{2}} e^{-p} \sqrt{2\pi} \left(1 + \frac{1}{12} a + \dots \right)$$

Nach dem Vorgange von Lagrange gibt Laplace**) die Eulersche Summationsformel durch den Beweis, dass

$$\sum y = \left[e^{h \frac{dy}{dx}} - 1 \right]^{-1} + \text{Const.}$$

wenn man in der Entwicklung der rechten Seite die Exponenten zugleich auf die Ordnung der Derivation $\frac{dy}{dx}$ bezieht und wenn $h \geq 1$ das Increment der unabhängigen Variablen x bedeutet. Es wird dann, wie man zeigen kann:

$$\sum y = \frac{1}{h} \int y dx - \frac{1}{2} y + \frac{h B(1)}{2!} y' - \frac{h^3 B(2)}{4!} y''' \pm \dots + \text{Const.}$$

*) Die Integrale von der Form $\int_{-\infty}^{\infty} t^{2n+1} e^{-t^2} dt$ sind = 0.

**) V. Lacroix, Grand Traité, 2. édit. t. III, p. 98.

Berichtigungen.

Seite 126, 10. Zeile v. o. lies: $\frac{s+1}{s} < \frac{rs+r}{rs-s}$.

» 128, 14. » v. o. » 42787536.

» » 17. » v. o. » 44623980.

» » 13. » v. u. » 25500 Versuchen.

Benutzte Quellen und Werke.

- Bernoulli J.*, Quaestiones nonnullae de usuris cum solutione problematis de sorte Aleatorum in Ephemerid. Gallic. Act. erud. Lips. 1690, p. 219.
- Bernoulli J.*, Ars conjectandi, Basileae 1713.
- Montmort*, Essai d'analyse sur le jeu de hasard. Paris 1708.
- Moivre A.*, The doctrine of chances or a method of calculating probability of events in play. London 1738.
- Miscellanea analytica de seriebus et quadraturis. London 1730.
- Miscellaneis analyticis supplementum.
- Stirling*, Methodus differentialis sive tractatus de summatione et interpolatione serierum infinitarum. London 1730.
- Simpson*, Treatise on the nature and laws of chance. London 1740.
- MacLaurin*, Treatise of Fluxions. Edinburgh 1742.
- Bayes*, An Essay towards solving a Problem in the doctrine of Chances. Phil. Transact. 1763, p. 370.
- Euler, L.*, Institutiones calculi differentialis. Petropoli 1755.
- Lagrange*, Mémoire sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations. Miscell. Taur. Tom. V, Turin 1775.
- Laplace*, Mémoire sur les probabilités. Mémoires de l'académie royale des sciences de Paris 1778, 1782.
- Essai philosophique sur les probabilités. 3. éd. Paris 1812.
- Théorie analytique des probabilités. Paris 1812.
- Lacroix*, Traité du calcul différentiel et du calcul intégral. t. s. Paris 1797.. (7. éd. par Hermite et Serret.)
- Poisson*, Recherches sur la probabilité des jugements en matière criminelle et en matière civile. Paris 1837.
- Liagre*, Calcul des probabilités. Bruxelles 1852.
- Serret*, Comptes rendus hebdomadaires des séances de l'académie des sciences, année 1860, t. I, p. 1662: Mémoire sur l'évaluation approchée du produit $1 \cdot 2 \cdot 3 \cdot \dots \cdot x$, lorsque x est un très grand nombre, et sur la formule de Stirling.
- Todhunter*, History of the mathematical theorie of probability. Cambridge und London 1865.
- Hagen, G.*, Grundzüge der Wahrscheinlichkeitsrechnung. 3. Aufl. Berlin 1882.
- Meyer*, Vorlesungen über Wahrscheinlichkeitsrechnung. Deutsche Ausgabe v. E. Czuber. Leipzig 1879.
-