mit $H_J := \inf_{x\in J} H(x|p)$. Für $J=\{1\}$ ist dies die Situation bei den 1-Runs. Für den allgemeineren Fall nutzt man im Beweis (3). In der Fachliteratur wird ein Gesetz wie für $R_n$ oder $R_n^J$ *Erdös-Rényi-Gesetz* genannt. Sie bekommen zunehmend Bedeutung in Anwendungen wie zum Beispiel der mathematischen Biologie [4].

## Literatur


[1] Eichelsbacher, P., (1999): Wie gut ist die Tschebychev-Ungleichung? Mathematik in der Schule.

[2] Erdös, P. and Rényi, A. (1970): On a new law of large numbers, Journ. Analyse Mathématique, 23, 103-111.

[3] Krengel, U. (1991): Einführung in die Wahrscheinlichkeitstheorie und Statistik, Vieweg Studiem, Aufbaukurs Mathematik, Braunschweig, 3. erweiterte Auflage.

[4] Waterman, M. (1995): Introduction to computational biology, Capman & Hill.



Anschrift des Verfassers
Peter Eichelsbacher
Fakultät für Mathematik
Ruhr-Universität Bochum
NA 3 / 68
44780 Bochum
peter.eichelsbacher@ruhr-uni-bochum.de


# Eine Formel von de Moivre


PETER EICHELSBACHER, BOCHUM



***Zusammenfassung:*** *Wir diskutieren eine in der Literatur fast in Vergessenheit geratene Formel von de Moivre zur Berechnung des absoluten Fehlers zwischen der relativen Häufigkeit und dem Erwartungswert in einer Bernoulli–Kette. Dabei geben wir eine historische Übersicht über die Entwicklung dieser Formel Bezug nehmend auf einen Artikel von Persi Diaconis und Sandy Zabell in Statistical Science.*


## 1 Einleitung

Das sicher am häufigsten im Stochastikunterricht behandelte Zufallsexperiment ist das des $n$-maligen unabhängigen Münzwurfs. Gegeben sei $\Omega=\{E,M\}$, wobei $E$ für Erfolg und $M$ für Mißerfolg steht. Die Wahrscheinlichkeit für Erfolg sei $0<p<1$ (der Stochastiker interessiert sich nicht für den Fall $p=0$ oder $p=1$). Dann setze $\Omega_n=\{E,M\}^n$, die Menge der $E$-$M$-Folgen der Länge $n$. Die Wahrscheinlichkeit für jedes $\omega=(\omega_1,\dots,\omega_n)\in\Omega_n$ sei gegeben durch $p(\omega)=p^k(1-p)^{n-k}$, wobei $k$ die Anzahl der $E$'s in der Folge $\omega_1,\dots,\omega_n$ bezeichnet. $X$ sei die Anzahl der Erfolge in der Bernoulli-Kette der Länge $n$ zum Parameter $p$. Darunter versteht man nichts anderes als den gerade beschriebenen Wahrscheinlichkeitsraum $(\Omega_n,p)$. Setze $X_i=1$, falls $\omega_i=1$ und $X_i=0$ sonst, $1\le i\le n$, so folgt $X=X_1+X_2+\dots+X_n$. Es gilt

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}=:b(k;n,p)$$

für $k\in\{0,\dots,n\}$. Die Zufallsgröße mit obiger Verteilung heißt binomialverteilt mit Parametern $n$ und $p$. Es gilt $\mathbb{E}(X)=np$ und $V(X)=np(1-p)$, wobei $\mathbb{E}(X)$ den Erwartungswert von $X$ bezeichnet und $V(X)$ die Varianz von $X$. Eine gute Darstellung findet man zum Beispiel im Buch von Barth und Haller [1]. Mit diesem Modell der Bernoulli-Kette arbeitet man auch in der *Statistik*, wo es darum geht, aus bekanntem $n$ und einer Beobachtung von sagen wir $k$ Erfolgen Rückschlüsse auf den unbekannten Parameter $p$ zu ziehen. Nun kann man Schülern schnell klar machen, dass die sogenannte *naive Schätzung* $k/n$ (in der Sprache der Zufallsgrößen $X/n$) ein vernünftiger Schätzer für $p$ ist. Es gilt

$$\mathbb{E}(X/n)=np/n=p.$$

Im Mittel kommt also bei dieser Schätzung der wahre Parameter raus (man nennt das in der Fachsprache *Erwartungstreue* des Schätzers $X/n$). Die Erwartungstreue ist quasi ein erstes Indiz für die Tauglichkeit der Wahl $X/n$. Weiter kann man plausibel machen, dass der Fehler dieser Schätzung in der Form $X/n-p$ dargestellt werden sollte. Da aber $X/n-p$ eine Zufallsgröße ist, kann man den mittleren Fehler bestimmen, nun gilt aber bekanntlich $\mathbb{E}(X/n-p)=\mathbb{E}(X/n)-p=p-p=0$. Das einfache Subtrahieren liefert also immer einen "Nullfehler". Somit ist $X/n-p$ kein geieigneter Kandidat für eine Maßzahl des Fehlers unserer Schätzung. Was ist nun ein nahe liegender Abstand zwischen zwei Zahlen? Vermutlich kommt fast jeder Schüler zu der Aussage der *absoluten Differenz*. Und so ist

$$D_n(p) := \mathbb{E}(|X/n - p|)$$

der *mittlere absolute Fehler* der Schätzung. Interessanterweise verfolgt man diesen Ausdruck im Bereich der Schule, aber auch im Hochschulbereich nicht (oder nur sehr selten). Meist schließt man mit dem Kommentar, dass es sich mit Absolutbeträgen schwer rechnet, weil die Definition von $|\cdot|$ über eine Fallunterscheidung erfolgt. Schnell geht man zum Begriff des *mittleren quadratischen Fehlers* über, in unserem Zusammenhang definiert als

$$S_n(p) := \sqrt{\mathbb{E}((X/n - p)^2)}.$$

Man bemerkt, dass dieser Ausdruck mehr oder weniger das gleiche wie $D_n(p)$ misst, und kann natürlich sehr schnell eine geschlossene Formel präsentieren: Wir wissen bereits, dass $V(X) = n\,p\,(1-p)$ ist, somit ist $S_n(p)^2 = V(X/n) = V(X)/n^2 = p\,(1-p)/n$. Also

$$S_n(p) = \sqrt{\frac{p(1-p)}{n}}.$$

Ist es didaktisch aber wirklich vertretbar, den eher natürlicheren Begriff $D_n(p)$ durch die quadratische Abweichung zu ersetzen? Bei einem Test in der Klasse ergibt sich sicher eher das Bild, dass der Begriff der mittleren quadratischen Abweichung ungewöhnlich, weniger intuitiv ist. Wie kommt man auf das Quadrat? Nun, es gibt natürlich neben der schnellen Berechenbarkeit von $S_n(p)$ noch weitere Vorteile dieses Begriffs, die man aber in der Regel im Schulunterricht nicht wirklich ausarbeiten kann. Das Stichwort ist die *Methode der kleinsten Quadrate* von Gauß und das in den Anwendungen damit verbundene Verfahren der *linearen Regression*.

Wir wollen in diesem Aufsatz zunächst rein mathematische Vorteile von $D_n(p)$ gegenüber $S_n(p)$ beschreiben. Die gibt es durchaus. Dies soll Motivation genug sein, sich dann der Berechnung von $D_n(p)$ zu widmen, im Fall der Bernoulli-Kette. Es ergibt sich ein interessanter historischer Ausflug bis zurück zu Abraham de Moivre.

## 2 Vorteile des absoluten mittleren Fehlers

Schauen wir uns einmal einen mathematischen Vergleich von $D_n(p)$ und $S_n(p)$ an. Nach einer wichtigen Formel für die Varianz einer Zufallsvariablen $X$ gilt $V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ (man nennt dies manchmal die Identität von Steiner). Wenden wir diese Identität auf die Zufalls größe $|X/n - p|$ an, so folgt

$$V\left(\left|\frac{X}{n} - p\right|\right) = \mathbb{E}\left(\left|\frac{X}{n} - p\right|^2\right) - \left(\mathbb{E}\left(\left|\frac{X}{n} - p\right|\right)\right)^2 \geq 0,$$

denn die Varianz ist per Definition nichtnegativ. Daraus folgt aber sofort

$$\mathbb{E}\left(\left|\frac{X}{n} - p\right|^2\right) \geq \left(\mathbb{E}\left(\left|\frac{X}{n} - p\right|\right)\right)^2$$

und, wenn man die Wurzel zieht, sofort

$$S_n(p) \geq D_n(p).$$

Die erste Beobachtung in Worten: Der mittlere quadratische Fehler ist immer größer als der mittlere absolute Fehler. $D_n(p)$ beschreibt in diesem Sinne also den mittleren Fehler besser. Ein weiteres Maß für die Qualität eines Schätzers ist der *Variationskoeffizient*, die relative, d.h. auf den Erwartungswert $p$ (den zu schätzenden Wert) bezogene Standardabweichung: Bei Wahl von $S_n(p)$ ist

$$\frac{S_n(p)}{\mathbb{E}(X/n)} = \sqrt{\frac{p(1-p)}{np^2}} = \sqrt{\frac{(1-p)}{n\,p}}$$

gemeint. Dieser Wert verhält sich für kleine Werte von $p$ ungünstig:

$$\lim_{p\to 0} \frac{S_n(p)}{\mathbb{E}(X/n)} = \infty.$$

Der Quotient strebt von der Größenordnung $p^{-1/2}$ gegen $\infty$. Anschaulich bedeutet dies, das kleine Häufigkeiten schwerer zu schätzen sind. Um diese Probleme zu vermeiden, kann ebenfalls besser $D_n(p)/p$ herangezogen werden, da

$$\lim_{p\to 0} \frac{D_n(p)}{p} = 2$$

gilt. Wir zeigen diese Formel allerdings erst im Rahmen der Ausführungen der angekündigten Formel von de Moivre.

Das Fazit ist: Es gibt rein mathematische Vorteile des Abstandsbegriffs $D_n(p)$, die es zumindest erklären, diesen Begriff durchaus – entgegen des Trends in den meisten Lehrbüchern – etwas intensiver zu betrachten. Wäre die Berechenbarkeit von $D_n(p)$ aber dennoch ein großes Problem, würde die Diskussion in diesem Abschnitt trotzdem nicht viel Gewinn bringen. Aber man kann $D_n(p)$ erstaunlich einfach berechnen.

## 3 Die Formel von de Moivre

Bereits 1730 gab Abraham de Moivre in [4] die folgende Formel an, die wir in Abschnitt 4 beweisen werden.

**Satz 1**

$$\mathbb{E}\big(|X-n\,p|\big) = 2\,v\,(1-p)\,b(v;n,p),$$

*wobei* $v$ *die eindeutig bestimmte ganze Zahl mit* $n\,p < v \leq n\,p+1$ *ist.*

Man nennt diese Formel auch *geschlossene Form*-Darstellung des mittleren absoluten Fehlers in der Bernoulli-Kette. Geschlossen bedeutet hier, dass die Formel nicht in Form von Fallunterscheidungen angegeben werden muss. Der auftretende Absolutbetrag suggeriert doch, dass Ausdrücke der Form $\sum_{k\leq n\,p} b(k;n,p)$ vorkommen. Wenn die Formel in Satz 1 durch $n$ dividiert wird, ergibt sich

$$D_n(p) = \frac{2\,v\,(1-p)}{n}\,b(v;n,p).$$

Daraus können wir schnell das folgende bereits angekündigte Resultat ableiten:

**Satz 2** *In der obigen Notation gilt*

$$\lim_{p\to 0} \frac{D_n(p)}{p} = 2.$$

*Beweis:* Da wir $p$ klein werden lassen, können wir zu festem $n \in \mathbb{N}$ $p$ so klein wählen, dass $n\,p < 1$ gilt. Dann ist in Satz 1 das durch $n\,p < v \leq n\,p+1$ eindeutig bestimmte $v$ gleich 1 (und es bleibt 1, auch wenn wir $p$ noch kleiner wählen und schließlich gegen 0 laufen lassen). Damit vereinfacht sich die Formel für $D_n(p)$ für alle $p$ mit $n\,p < 1$ zu $D_n(p) = 2(1-p)\,p(1-p)^{n-1}$, also $D_n(p)/p = 2(1-p)^n$ und dies konvergiert gegen 2 für $p \to 0$. Damit ist der Satz bewiesen.

Nutzen wir die Tatsache, dass der Wert $b(v;n,p)$ für jeden festen Wert $v$ für $n\to\infty$ gegen null konvergiert, so folgt aus de Moivres Formel auch unmittelbar

$$\lim_{n\to\infty} \mathbb{E}\left(\left|\frac{X}{n} - p\right|\right) = 0. \qquad (3)$$

Dies ist eine Form des Gesetzes der großen Zahlen, die erstmals de Moivre so formulierte. Historisch interessant ist die Bemerkung, dass de Moivre die Formel (3) zum Anlass nahm, die Approximation der Binomialverteilung näher zu studieren. Dies mündete dann in dem von ihm gegebenen Beweis des zentralen Grenzwertsatzes für den Fall $p = 1/2$, heute Satz von de Moivre und Laplace genannt. Er studiert die Approximation der Binomialverteilung durch die Gaußsche Normalverteilung. Dies verfolgen wir hier nicht weiter.

## 4 Eine Formel von Todhunter

Abraham de Moivre (1667–1754) schrieb das Buch *The Doctrine of Chances* [3] im Jahre 1718. Es entstanden zwei weitere Ausgaben in den Jahren 1738 und 1756. In der ersten Ausgabe findet man zu der in Satz 1 gegebenen Formel keinen Hinweis. Erst in den weiteren beiden Ausgaben wird die Formel in einem wahrscheinlichkeitstheoretischen Kontext vorgestellt und bewiesen. In der Ausgabe 1738 sind dies die Probleme 86 und 87, in der Ausgabe 1756 sind es die Probleme 72 und 73. Tatsächlich stellt de Moivre in einem weniger bekannten Werk mit dem Titel *Miscellanea Analytica* im Jahre 1730 die Frage nach der Berechnung des absoluten mittleren Fehlers $D_n(p)$ vor. Sie entstammt einer Frage von Sir Cuming aus dem Jahre 1721. De Moivre beweist in dieser Arbeit die Formel aus Satz 1 für den *symmetrischen* Fall $p = 1/2$. Er gibt in dieser Arbeit an, wie die Formel im asymmetrischen Fall $p \neq 1/2$ aussieht. Er gibt aber keinen Beweis!

Isaac Todhunter gibt im Jahre 1865 in seinem Werk mit dem Titel *A History of the Mathematical Theory of Probability, Seiten 182-183*, [9] einen Beweis im asymmetrischen Fall. Seine Formel liest sich in heutiger Sprache so:

**Satz 4** *Für alle natürlichen Zahlen* $0 \leq \alpha < \beta \leq n$ *gilt*

$$\sum_{k=\alpha}^{\beta} (k-n\,p)b(k;n,p) = \alpha\,(1-p)\,b(\alpha;n,p) - (n-\beta)\,p\,b(\beta;n,p).$$

Bevor der sehr elementare Beweis von Satz 4 gegeben wird, wollen wir schauen, wie daraus Satz 1 abgeleitet werden kann. Für alle $k \geq v$ ist nach Definition von $v$ $|k-n\,p| = k-n\,p$ und für alle anderen Werte gilt $|k-n\,p| = -(k-n\,p)$. Damit erhalten wir

$$\mathbb{E}(|X-n\,p|) = \sum_{k=v}^{n} (k-n\,p)b(k;n,p) - \sum_{k=0}^{v-1} (k-n\,p)b(k;n,p).$$

Wir wenden nun auf beide Summen den Satz 4 von Todhunter an:

$$\mathbb{E}(|X - np|) = v(1-p)b(v;n,p) + (n-(v-1))pb(v-1;n,p).$$

Der erste Summand ist in der gewünschten Form. Den zweiten Summanden formen wir um. Das Verhältnis zweier aufeinanderfolgender Binomialwahrscheinlichkeiten berechnet sich bekanntlich wie folgt:

$$\frac{b(k+1;n,p)}{b(k;n,p)} = \frac{(n-k)\,p}{(1-p)(k+1)}. \tag{5}$$

Daraus folgt unmittelbar

$$b(k+1;n,p)\,(1-p)\,(k+1) = (n-k)\,p\,b(k;n,p). \tag{6}$$

Setzen wir hierin $k = v-1$, so erhalten wir

$$(n-(v-1))\,p\,b(v-1;n,p) = v\,(1-p)\,b(v;n,p),$$

und damit insgesamt

$$\mathbb{E}(|X-np|) = 2\,v\,(1-p)\,b(v;n,p).$$

Dies ist die Formel von de Moivre.

Nun beweisen wir die Formel von Todhunter: Es gilt

$$\begin{aligned}\sum_{k=\alpha}^{\beta}(k-np)\,b(k;n,p) &= \sum_{k=\alpha}^{\beta}\big(k(1-p)-(n-k)p\big)\,b(k;n,p)\\ &= \sum_{k=\alpha}^{\beta}k(1-p)\,b(k;n,p) - \sum_{k=\alpha}^{\beta}(n-k)\,p\,b(k;n,p).\end{aligned} \tag{7}$$

Setzen wir (6) ein, so erhalten wir

$$\begin{aligned}\sum_{k=\alpha}^{\beta}(k-np)\,b(k;n,p) = {}&\\ &\sum_{k=\alpha}^{\beta}k(1-p)b(k;n,p)\\ &-\sum_{k=\alpha}^{\beta}(k+1)(1-p)b(k+1;n,p).\end{aligned}$$

Es bleibt also nur der erste Summand der ersten Summe und der letzte Summand der zweiten Summe übrig, und dies liefert Todhunters Formel.

Zugegeben, die Formel von de Moivre benötigt einen zwar elementaren aber sicher nicht offensichtlichen Beweis. Die Schönheit der Formel liegt - nach vollbrachtem Beweis - in der geschlossenen Form, die jede Fallunterscheidung entfernt. Diese steckt in unserem Beweis! Nach getaner Arbeit ist es also nicht mehr schwer, den mittleren absoluten Fehler in der Bernoulli-Kette zu bestimmen.

## 5 Abschließende Bemerkungen

Todhunter nahm in seinem Beweis übrigens an, dass $np$ eine natürliche Zahl ist. Er sah nicht, dass sein Beweis diese Annahme in Wirklichkeit nicht benötigt hätte. Die Formel von de Moivre bzw. die Formel von Todhunter wurde streng genommen in der hier vorliegenden Form erst von Poincaré 1896 in seinem Buch *Calcul des Probabilités* [8] bewiesen. Interessant ist, dass die hier vorgestellten Formeln immer wieder neu entdeckt wurden. So hat Frisch in [5] auf Seite 161 Todhunters Formel bewiesen und die Formel von de Moivre daraus abgeleitet. Frisch schien nicht gewußt zu haben, dass es die Formeln schon gab. Weitere “Neuentdeckungen” der Formeln finden sich 1930 in Gruder [6] und 1957 in Johnson [7]. Johnson betrachtet allerdings auch einige Verallgemeinerungen der genannten Formeln. Die Formel von de Moivre ist in den Werken der Vorgänger scheinbar immer so wenig hervorgehoben worden, dass Forscher in der späteren Schaffenszeit sie nicht entdeckten. Interessant ist, dass niemand ihren Zusammenhang mit dem Gesetz der großen Zahlen beschrieb. Die Formel blieb anscheinend für sich genommen eine kuriose Beobachtung.

Eine hübsche Anwendung von Satz 1 ist:

**Satz 8** *Für $\alpha > np$, $n \in \mathbb{N}$ und $p \in (0,1)$ gilt:*

$$\frac{\alpha}{n} \le \frac{1}{b(\alpha;n,p)}\sum_{k=\alpha}^{n} b(k;n,p) \le \frac{\alpha(1-p)}{\alpha - np}.$$

Der Beweis geht so: Sicher gilt

$$\alpha\sum_{k=\alpha}^{n} b(k;n,p) \le \sum_{k=\alpha}^{n} k\,b(k;n,p).$$

Auf der rechten Seite wenden wir nun Todhunters Formel an:

$$\sum_{k=\alpha}^{n} k\,b(k;n,p) = np\sum_{k=\alpha}^{n} k\,b(k;n,p) + \alpha\,(1-p)\,b(\alpha;n,p).$$

Daraus folgt die obere Schranke. Die untere Schranke erhält man einfacher:

$$n\sum_{k=\alpha}^{n} b(k;n,p) \ge \sum_{k=\alpha}^{n} k\,b(k;n,p) \ge \alpha b(\alpha;n,p).$$

Es stellt sich die Frage, ob es für andere Verteilungen der Stochastik auch eine geschlossene Form für die Bestimmung des mittleren absoluten Fehlers gibt. Bisher war alles auf die Binomialverteilung konzentriert. In der Arbeit von Diaconis und Zabell [2] findet man dies im Rahmen der Theorie orthogonaler Polynome vorgestellt. So sieht man zum Beispiel, dass für den Fall einer Standard-Normalverteilung gilt:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|x|\,e^{-x^2/2}\,dx=\frac{2}{\sqrt{2\pi}}.$$

Dies kann man natürlich auch direkt mittels der Integrationstheorie herleiten. Im Falle einer Poisson verteilten Zufallsgröße mit Parameter $\lambda$ ergibt sich:

$$\sum_{k=0}^{\infty}|k-\lambda|\,\frac{e^{\lambda}\lambda^{k}}{k!}=2\lambda\frac{e^{\lambda}\lambda^{[\lambda]}}{[\lambda]!},$$

wobei $[\lambda]$ die untere Gauß-Klammer bezeichnet, also die größte ganze Zahl kleiner oder gleich $\lambda$. Wir leiten hier diese beiden Formeln nicht her.

Abschließend geben wir noch einen kleinen Einblick in die Eigenschaften der Funktion $p\mapsto D_n(p)$. Dies ist eine hübsche Übung aus der Analysis. Wir betrachten das Intervall $(k-1)/n<p<k/n$ für $k=1,\ldots,n$. Das $\nu$ der Formel von de Moivre hat dann den Wert $k$, so dass sich $D_n(p)$ mit Hilfe dieser Formel darstellen lässt:

$$D_n(p)=2\binom{n-1}{k-1}p^k(1-p)^{n-k+1}.$$

Man kann $D_n$ hier zweimal differenzieren (Anwendung der Produktregel) und sieht dann: $D_n'(p)$ ist positiv für $p<k/(n+1)$, gleich null für $p=k/(n+1)$ und negativ für $p>k(n+1)$. Außerdem ist $D_n''(p)$ negativ. Also hat $D_n$ lokale Minima an den Stellen $p=1/n,2/n,\ldots,(n-1)/n$ und lokale Maxima an den Stellen $p=1/(n+1),2/(n+1),\ldots,n/(n+1)$ und ist in allen Zwischenbereichen eine konkave Funktion! Etwas schwerer zu berechnen ist dann die weitere Beobachtung, dass die lokalen Maxima der Reihe nach größer werden bis zum Wert $p=1/2$ und danach der Reihe nach kleiner werden. Das globale Maximum ist daher für ungerade $n$ an der Stelle $p=1/2$ gegeben, für gerade $n$ an den Stellen $p=1/2\pm 1/\{2(n+1)\}$. Diese Analyse ist etwas komplizierter und soll dem Leser als Übung bleiben.

Die Formeln von Todhunter und de Moivre zur Bestimmung des mittleren absoluten Fehlers im Falle eines binomialverteilten Modells halte ich für eine interessante Bereicherung des Mathematikunterrichts der Oberstufen. Die historische Entwicklung dieser eher nicht sehr berühmten Formel ist weiter ein Anreiz für das Studium von Originalquellen, hier besonders von [4].

## Literatur


[1] Barth, F. und Haller, R., *Stochastik, Leistungskurs*, Ehrenwirth Verlag, 1991.

[2] Diaconis, P. and Zabell, S., Closed form summation for classical distributions: variations on a theme of de Moivre, *Statistical Science. A Review Journal of the Institute of Mathematical Statistics*, Vol.6, No. 3, 284–302, 1991.

[3] De Moivre, A., *The Doctrine of Chances*, 1st ed. A. Millar, London, 1718. (2nd ed. 1738; 3rd ed. 1756)

[4] De Moivre, A., *Miscellanea Analytica de seriebus et Quadraturis*, J. Tonson and J. Watts, London, 1730.

[5] Frisch, R., *Solution d'un problém du calcul des probabilités*, Skandinavisk Aktuarietidskrift 7, 153–174, 1924.

[6] Gruder, O., *9th International Congress of Actuaries*, 2, 222, 1930.

[7] Johnson, N. L., *A note on the mean deviation of the binomial distribution*, Biometrika, 44, 532–533, 1957.

[8] Poincaré, H., *Calcul des Probabilités*, 1st ed. Georges Carré, Paris, 1896 (2nd ed. 1912, Gauthier-Villars, Paris)

[9] Todhunter, I., *A History of the Mathematical Theory of Probability*, Macmillan, London, 1865.



Anschrift des Verfassers<br>
Peter Eichelsbacher<br>
Fakultät für Mathematik<br>
Ruhr-Universität Bochum<br>
NA 3 / 68<br>
44780 Bochum<br>
peter.eichelsbacher@ruhr-uni-bochum.de