ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIV, 4 (1975)

T. GERSTENKORN (ŁÓDŹ)

BEMERKUNGEN ÜBER DIE ZENTRALEN UNVOLLSTÄNDIGEN UND ABSOLUTEN MOMENTE DER PÓLYA-VERTEILUNG (1)

1. Einleitung. Der Begriff der unvollständigen Momente ist im Jahre 1925 von Ragnar Frisch für die Binomialverteilung angewandt worden. Seit jener Zeit wurde diese Charakteristik von vielen Verfassern für verschiedene Zwecke, auf verschiedene zufällige Veränderliche angewandt und auf verschiedene Weise benutzt. Eine umfangreiche Übersicht der Entwicklung dieses Problems ist von dem Verfasser in der Monographie [4] aus dem Jahre 1971 gegeben worden. In der hier vorliegenden Arbeit ist eine Methode dargestellt. mit der sich die rechtsseitigen unvollständigen Zentralmomente der Pólya-Verteilung leicht entwickeln lassen. Dabei wird der von A. R. Kamat im Jahre 1963 gegebene Satz benutzt.

Der Paragraph 2 ruft kurz die bekannten Definitionen der verschiedenen Momente in Erinnerung. Im Paragraph 3 wird der Satz von Kamat zitiert. Im Paragraph 4 wird bewiesen, daß die Pólya-Verteilung die Voraussetzungen des Satzes von Kamat erfüllt und in den Abschnitten 4.4 und 4.5 werden die entsprechenden Rekursionsformeln für die Momente dieser Verteilung entwickelt und als Beispiel die vier ersten Momente ausgerechnet. Paragraph 5 erläutert die Spezialfälle, d.h. gibt die Momente für jene Verteilungen an, die auf der Pólya-Verteilung basieren. Paragraph 6 bringt die Anwendung der Ergebnisse der Paragraphen 4 und 5 mit einer Ermittlung der Formeln für die absoluten Zentralmomente und der mittleren Abweichung aller dargestellten Verteilungen.

2. Definitionen. Wir betrachten eine diskrete ganzzahlige zufällige Veränderliche X mit der Wahrscheinlichkeitsfunktion $p_k = \mathrm{P}(X = k)$. Der Ausdruck

$$\mu_r'(s) = \sum_{k=s}^{\infty} (k-c)^r p_k, \qquad r = 1, 2, 3, \ldots, \tag{2.1}$$

---

(1) Die Arbeit entstand während eines Studienaufenthalts an der Universität Bonn (Institut für Landw. Botanik — Biometrie, Prof. Dr. F. Weiling). Der Verfasser bedankt sich bei dem Deutschen Akademischen Austauschdienst für die Gewährung eines Aufenthaltsstipendiums.

wobei s eine beliebige Zahl aus dem Wertebereich von X ist, wird das *rechtsseitige unvollständige Moment* der r-ten Ordnung in Bezug auf eine beliebige Konstante c genannt.

Wenn $s = -\infty$ ist, dann ergibt die Formel (2.1) ganz einfach das *vollständige Moment* μ_r' in Bezug auf eine Konstante c:

$$\mu_r' = \sum_{k=-\infty}^{\infty} (k-c)^r p_k, \qquad r = 1, 2, 3, \ldots$$

Wenn $c = \mathrm{E}(X) = a$ den Erwartungswert darstellt, bekommen wir daraus das *rechtsseitige unvollständige Zentralmoment* $\mu_r(s)$,

$$\mu_r(s) = \sum_{k=s}^{\infty} (k-a)^r p_k, \qquad r = 1, 2, 3, \ldots, \tag{2.2}$$

und im Fall $s = -\infty$ ohne weiteres das *Zentralmoment* μ_r:

$$\mu_r = \sum_{k=-\infty}^{\infty} (k-a)^r p_k, \qquad r = 2, 3, \ldots \tag{2.3}$$

Es ist unmittelbar zu sehen, daß, wenn $a = 0$ ist, die Formeln (2.2) und (2.3) die Gestalt der gewöhnlichen Momente $\alpha_r(s)$ oder α_r annehmen.

3. Der Satz von A. R. Kamat. Es nehme die zufällige Veränderliche X die Werte $m, m+1, \ldots, n$ an. Es sei $p_k = \mathrm{P}(X = k)$ ihre Wahrscheinlichkeitsfunktion, die nur von den Parametern $\theta = (\theta_1, \theta_2, \ldots, \theta_l)$ abhängig sein kann. Es sei weiter vorausgesetzt, daß die Konstante c auch nur von den Parametern θ abhängt, d.h. $c = c(\theta)$, und daß

$$(k-c)p_k = U_k - U_{k+1}, \tag{A}$$

$$U_{k+1} = p_k\big(A_1(k-c)^2 + A_2(k-c) + A_3\big), \tag{B}$$

wobei A_1, A_2, A_3 nur von den eingangs zitierten Parametern abhängig sind und $(r-1)A_1 \neq 1$ für $r = 2, 3, \ldots$, sowie

$$U_{n+1} = 0 \tag{C}$$

ist.

Aus diesen Voraussetzungen ergibt sich die folgende Relation:

$$\begin{aligned}\mu_r'(s) &= \big(1-(r-1)A_1\big)^{-1} \times \\ &\times \Big\{(s-c)^{r-1} U_s + (r-1)\big(A_2\mu_{r-1}'(s) + A_3\mu_{r-2}'(s)\big) + \\ &+ \sum_{i=0}^{r-3} \binom{r-1}{i} \big(A_1\mu_{i+2}'(s) + A_2\mu_{i+1}'(s) + A_3\mu_i'(s)\big)\Big\}.\end{aligned} \tag{3.1}$$

Es sei bemerkt, daß diese Formel folgendermaßen dargestellt werden kann:

$$(3.2)\qquad \mu_r'(s) = \big(1-(r-1)A_1\big)^{-1}\Big\{(s-c)^{r-1}U_s + A_1\sum_{i=0}^{r-3}\binom{r-1}{i}\mu_{i+2}'(s) + \\ + \sum_{i=0}^{r-2}\binom{r-1}{i}\big(A_2\mu_{i+1}'(s)+A_3\mu_i'(s)\big)\Big\}.$$

4. Rekursionsformeln für die Momente der Pólya-Verteilung. Wir werden beweisen, daß eine Pólya-verteilte zufällige Veränderliche die oben gegebenen Bedingungen erfüllt, wenn für c der Erwartungswert $c = \mathrm{E}(X) = np$ angenommen wird. Mit Hilfe dieses Ergebnisses können wir in den Abschnitten 4.4 und 4.5 die entsprechenden Rekursionsformeln für die unvollständigen Zentralmomente entwickeln.

4.1. Der Ausdruck

$$x^{[k,h]} = x(x-h)(x-2h)\ldots\big(x-(k-1)h\big)$$

wird ein *faktorielles Polynom* der Ordnung k mit dem Schritt h genannt.

Bedient man sich dieser Bezeichnung, so läßt sich die Wahrscheinlichkeitsfunktion der Pólya-Verteilung in der Form

$$(4.1)\qquad p_k = \mathrm{P}(X=k) = \binom{n}{k}\frac{p^{[k,-a]}q^{[n-k,-a]}}{1^{[n,-a]}}$$

schreiben, wobei $0 < p < 1$, $q = 1-p$ ist, und die Zahlen k und a nachstehende Bedingungen erfüllen: $k = 0, 1, 2, \ldots, n$, $-ka \leqslant p$ und $-(n-k)a \leqslant q$.

Für die Wahrscheinlichkeitsfunktion (4.1) ergibt sich die folgende Relation:

$$(4.2)\qquad p_k = \frac{\big(n-(k-1)\big)\big(p+(k-1)a\big)}{k\big(q+(n-k)a\big)}p_{k-1}$$

(siehe [4], S. 34, (3.34)).

Zum Beweis, daß die Verteilung (4.1) die Bedingung (A) des zitierten Satzes erfüllt, formen wir die Differenz $k-np$ folgendermaßen um:

$$k-np = k\big(q+(n-k)a\big)-(n-k)(p+ka)$$

(siehe [4], S. 34, (3.1)).

Wenn wir jetzt (4.2) anwenden, bekommen wir

$$(k-np)p_k = (n-k+1)\big(p+(k-1)a\big)p_{k-1}-(n-k)(p+ka)p_k.$$

Bedienen wir uns der Bezeichnung

$$(4.3)\qquad U_{k+1} = (n-k)(p+ka)p_k,$$

so dürfen wir

$$(4.4) \qquad (k-np)p_k = U_k - U_{k+1}$$

schreiben, was mit der Erfüllung der Bedingung (A) gleichwertig ist.

4.2. Wir kommen jetzt zum Beweis, daß die Verteilung (4.1) die Bedingung (B) erfüllt. Zu diesem Zweck stellen wir das Produkt $(n-k)(p+ka)$ in Gestalt der in Bezug auf die Potenzen von $k-np$ entwickelten Summe dar. Wir haben dann

$$(n-k)(p+ka) = npq(1+na)-(p-nqa+npa)(k-np)-a(k-np)^2$$

(siehe [4], S. 34, (3.2)).

Man sieht, daß die Bedingung (B) erfüllt ist, wenn wir

$$(4.5) \qquad A_1 = -a, \qquad A_2 = na(q-p)-p, \qquad A_3 = npq(1+na)$$

und $c = np$ annehmen. Wir setzen dabei voraus, daß $-(r-1)a \neq 1$ ist.

4.3. Wir bemerken jetzt, daß sich zum Beweis der Formel (3.1) bei Benutzung von (4.4) $\mu_r'(s)$ folgendermaßen darstellen läßt:

$$\begin{aligned}\mu_r'(s) &= \sum_{k=s}^{n}(k-c)p_k(k-c)^{r-1} = \sum_{k=s}^{n}(k-c)^{r-1}(U_k - U_{k+1}) \\ &= \sum_{k=s}^{n}[(k-c)^{r-1}U_k-(k+1-c)^{r-1}U_{k+1}]+ \\ &\quad + \sum_{k=s}^{n}U_{k+1}[(k+1-c)^{r-1}-(k-c)^{r-1}].\end{aligned}$$

Der erste Bestandteil dieser Summe ergibt

$$(4.6) \qquad (s-c)^{r-1}U_s-(n+1-c)^{r-1}U_{n+1},$$

was nach Berücksichtigung von (C)

$$(4.7) \qquad (s-c)^{r-1}U_s$$

erreicht wird.

Es sei noch bemerkt, daß sich nichts an diesem Ergebnis ändert, wenn n durch $n+1$ ersetzt wird. Wir haben nämlich $p_{n+1} = 0$, also

$$\mu_r'(s) = \sum_{k=s}^{n}(k-c)^r p_k = \sum_{k=s}^{n+1}(k-c)p_k$$

und daraus bekommen wir statt (4.6) die folgende Relation:

$$(s-c)^{r-1}U_s-(n+2-c)^{r-1}U_{n+2}.$$

Wir sehen jedoch, daß

$$U_{n+2} = p_{n+1}\big(A_1(k-c)^2 + A_2(k-c) + A_3\big) = 0$$

ist; also ohne Unterschied gelangt man zu (4.7). Man sieht klar, daß die Bedingung (C) durch die Bedingung $U_{n+2} = 0$ ersetzt werden kann und daß diese Bedingung im Fall der Pólya-Verteilung erfüllt ist.

4.4. Es sind also alle Voraussetzungen des Satzes von Kamat erfüllt und so lassen sich nun die unvollständigen rechtsseitigen Momente der Pólya-Verteilung in Gestalt der Rekursionsformel (3.2) darstellen.

Nach Berücksichtigung von (4.2) und (4.3) haben wir

$$U_s = sp_s\big(q+(n-s)a\big). \tag{4.8}$$

Andererseits, wenn wir in (3.2) den Wert $r = 1$ annehmen, stellen wir fest, daß

$$\mu_1(s) = U_s \tag{4.9}$$

ist.

Setzen wir jetzt die erlangten Formeln (4.5), (4.8) und (4.9) wieder in (3.2) ein, dann bekommen wir endlich die Rekursionsformel für die rechtsseitigen unvollständigen Zentralmomente der Pólya-Verteilung:

$$\mu_r(s) = \big(1+(r-1)\,a\big)^{-1}\Big\{(s-np)^{r-1}\mu_1(s) - a\sum_{i=0}^{r-3}\binom{r-1}{i}\mu_{i+2}(s) + \sum_{i=0}^{r-2}\binom{r-1}{i}\Big[\big(na(q-p)-p\big)\mu_{i+1}(s) + npq(1+na)\mu_i(s)\Big]\Big\}. \tag{4.10}$$

Wenn s den Wert Null annimmt, dann ergibt die Formel (4.10) sofort die vollständigen Zentralmomente:

$$\mu_r = \big(1+(r-1)a\big)^{-1}\Big\{-a\sum_{i=0}^{r-3}\binom{r-1}{i}\mu_{i+2} + \sum_{i=0}^{r-2}\binom{r-1}{i}\Big[\big(na(q-p)-p\big)\mu_{i+1} + npq(1+na)\mu_i\Big]\Big\}. \tag{4.11}$$

Die Formeln (4.10) und (4.11) bilden eine andere Gestalt der Formeln (3.5) und (3.7), die in [4], S. 35-36, im Jahre 1971 mit einer anderen Methode bewiesen wurden. Die Formel (4.11) ist auch im Jahre 1972 von Mühlbach [10], S. 174, angegeben worden. (Vergl. das Zusammentreffen der angewandten Methode.)

Aus Relationen (4.8)-(4.11) ergeben sich die folgenden Formeln für die ersten vier Momente der Pólya-Verteilung:

$$\mu_1(s) = sp_s\big(q+(n-s)a\big) \tag{4.12}$$

(siehe auch [4], S. 35, (3.6)),

$$\mu_2(s) = (1+a)^{-1}\big\{\big(s-p(n+1)+na(q-p)\big)\mu_1(s)+npq(1+na)\mu_0(s)\big\},$$

$$\begin{aligned}\mu_3(s) = (1+2a)^{-1}\big\{&\big[2\big(na(q-p)-p\big)-a\big]\mu_2(s)+\\&+[(s-np)^2+na(q-p)-p+npq(1+na)]\mu_1(s)+\\&+npq(1+na)\mu_0(s)\big\},\end{aligned}$$

$$\begin{aligned}\mu_4(s) = (1+3a)^{-1}\big\{&3[na(q-p)-(p+a)]\mu_3(s)+\\&+\big[3\big(na(q-p)-p+npq(1+na)\big)-a\big]\mu_2(s)+\\&+[(s-np)^3+na(q-p)-p+3npq(1+na)]\mu_1(s)+\\&+npq(1+na)\mu_0(s)\big\}\end{aligned}$$

und

$$\mu_2 = (1+a)^{-1}npq(1+na),$$

$$\mu_3 = \big((1+a)(1+2a)\big)^{-1}npq(1+na)\big\{2\big(na(q-p)-p\big)+1\big\},$$

$$\begin{aligned}\mu_4 = \big((1+a)(1+2a)(1+3a)\big)^{-1}npq(1+na)\big\{&3[na(q-p)-(p+a)]\times\\&\times\big[2\big(na(q-p)-p\big)+1\big]+\\&+\big[3\big(na(q-p)-p+npq(1+na)\big)-a\big](1+2a)+(1+a)(1+2a)\big\}.\end{aligned}$$

4.5. Wenn wir die Wahrscheinlichkeit p, die in der Formel (4.1) auftritt, als jene Wahrscheinlichkeit interpretieren, aus einer Urne mit N Kugeln eine Kugel einer bestimmten Farbe (z.B. der weißen) zu ziehen, und wir annehmen, daß $a = S/N$ ist, wobei S die Zahl der zugegebenen oder entnommenen Kugeln nach dem Pólyaschen Schema ist, dann nimmt die Formel (4.1) folgende Gestalt an:

$$p_k = \mathrm{P}(X=k) = \binom{n}{k}\frac{(Np)^{[k,-S]}(Nq)^{[n-k,-S]}}{N^{[n,-S]}}, \tag{4.13}$$

wobei $q = 1-p$, n die Zahl der Ziehungen und k die Zahl der gezogenen (z.B. der weißen) Kugeln ist. Man sieht hierbei, daß die Zahl k mit der folgenden Ungleichung in Zusammenhang steht:

$$\max\left(0,\, n-1+\frac{Nq-1}{S}\right) \leqslant k \leqslant \min\left(n,\, \frac{1-Np}{S}+1\right).$$

In diesem Fall modifiziert sich die Formel (4.10) auf folgende Weise:

$$\mu_r(s) = \big(N + (r-1)S\big)^{-1}\Big\{N(s-np)^{r-1}\mu_1(s) - S\sum_{i=0}^{r-3}\binom{r-1}{i}\mu_{i+2}(s) + \sum_{i=0}^{r-2}\binom{r-1}{i}\Big[\big(nS(q-p)-Np\big)\mu_{i+1}(s) + npq(N+nS)\mu_i(s)\Big]\Big\}, \tag{4.14}$$

wobei

$$\mu_1(s) = \frac{s}{N}\big(Nq + (n-s)S\big)p_s \tag{4.15}$$

ist.

Die Formel (4.11) können wir nun folgendermaßen darstellen:

$$\mu_r = \big(N + (r-1)S\big)^{-1}\Big\{-S\sum_{i=0}^{r-3}\binom{r-1}{i}\mu_{i+2} + \sum_{i=0}^{r-2}\binom{r-1}{i}\Big[\big(nS(q-p)-Np\big)\mu_{i+1} + npq(N+nS)\mu_i\Big]\Big\}. \tag{4.16}$$

Die ersten vier Zentralmomente der Pólya-Verteilung haben in diesem Fall die folgende Gestalt:

$\mu_1(s)$ wie in (4.15),

$$\mu_2(s) = (N+S)^{-1}\Big\{\big[N\big(s-p(n+1)\big) + nS(q-p)\big]\mu_1(s) + npq(N+nS)\mu_0(s)\Big\},$$

$$\mu_3(s) = (N+2S)^{-1}\Big\{\big[2\big(nS(q-p)-Np\big)-S\big]\mu_2(s) + [N(s-np)^2 + nS(q-p) - Np + npq(N+nS)]\mu_1(s) + npq(N+nS)\mu_0(s)\Big\},$$

$$\mu_4(s) = (N+3S)^{-1}\Big\{3\,[nS(q-p)-Np-S]\mu_3(s) + \big[3\big(nS(q-p)-Np+npq(N+nS)\big)-S\big]\mu_2(s) + [N(s-np)^3 + nS(q-p) - Np + 3npq(N+nS)]\mu_1(s) + npq(N+nS)\mu_0(s)\Big\},$$

$$\mu_2 = (N+S)^{-1}npq(N+nS),$$

$$\mu_3 = \big((N+S)(N+2S)\big)^{-1}npq(N+nS)\big\{2\big(nS(q-p)-Np\big)+N\big\},$$

$$\mu_4 = \big((N+S)(N+2S)(N+3S)\big)^{-1}npq(N+nS)\Big\{3\,[nS(q-p)-Np-S]\times\big[2\big(nS(q-p)-Np\big)+N\big] + \big[3\big(nS(q-p)-Np+npq(N+nS)\big)-S\big](N+2S) + (N+S)(N+2S)\Big\}.$$

5. Spezialfälle.

5.1. Die Formeln (4.10), (4.11) oder (4.14), (4.16) können als Grundlage dienen, die entsprechenden Formeln für die Momente der folgenden Wahrscheinlichkeitsverteilungen zu erhalten: der binomialen, Poissonschen, hypergeometrischen, negativ-binomialen, Pólya-Eggenbergerschen, geometrischen und logarithmischen Verteilung.

Wenn wir $a = 0$ annehmen, so folgt aus (4.1) die Binomialverteilung

$$p_k = \mathrm{P}(X = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, 2, \ldots, n, \tag{5.1}$$

und aus (4.10) die Formel für die unvollständigen Zentralmomente

$$\mu_r(s) = (s - np)^{r-1}\mu_1(s) + p\sum_{i=0}^{r-2}\binom{r-1}{i}\big(nq\mu_i(s) - \mu_{i+1}(s)\big), \tag{5.2}$$

wobei, nach Berücksichtigung von (4.12),

$$\mu_1(s) = sqp_s \tag{5.3}$$

ist, während aus (4.11) die Formel für die vollständigen Zentralmomente

$$\mu_r = p\sum_{i=0}^{r-2}\binom{r-1}{i}(nq\mu_i - \mu_{i+1}) \tag{5.4}$$

folgt.

Stellen wir in Formel (5.1) die Limesbetrachtung

$$\lim_{n\to\infty} np = \lambda$$

an, so folgt die Poissonsche Verteilung mit

$$p_k = \mathrm{P}(X = k) \doteq \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots \tag{5.5}$$

Desgleichen erhält man aus der oben angegebenen Limesbetrachtung in (5.2) die Formel für die unvollständigen Zentralmomente der Verteilung (5.5),

$$\mu_r(s) = (s - \lambda)^{r-1}\mu_1(s) + \lambda\sum_{i=0}^{r-2}\binom{r-1}{i}\mu_i(s), \tag{5.6}$$

wobei, nach (5.3),

$$\mu_1(s) = sp_s \tag{5.7}$$

folgt, während wegen (5.4)

$$\mu_r = \lambda\sum_{i=0}^{r-2}\binom{r-1}{i}\mu_i \tag{5.8}$$

ist.

Die Begründung jener Betrachtung ist wie folgt. Für die Verteilung einer Zufallsvariable X, die nur endlich viele Werte annimmt (d.h. im besonderen Fall für die binomiale Verteilung) haben wir

$$\mu_r(s) = \sum_{k=s}^{n} (k-a)^r p_k$$

und nächst

$$\mu_r(s) = \sum_{k=s}^{\infty} (k-a)^r p_k,$$

weil $p_k = 0$ für $k > n$. Wenn $n \to \infty$ und $q_k = \mathrm{Q}(X = k)$ die Grenzverteilung für die Verteilung $\mathrm{P}_n(X = k) = p_k$ und $\mu_r^*(s)$ das Moment der Grenzverteilung sind, dann

$$\mu_r(s) \to \sum_{k=s}^{\infty} (k-a)^r q_k = \mu_r^*(s).$$

Die Voraussetzung $S = -1$ in (4.13) ergibt die hypergeometrische Verteilung

$$p_k = \mathrm{P}(X = k) = \binom{n}{k} \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}, \tag{5.9}$$

$$\max(0, n - Nq) \leqslant k \leqslant \min(n, Np),$$

und aus derselben Voraussetzung in (4.14)-(4.16) folgen die Formeln:

$$\mu_r(s) = (N-r+1)^{-1}\Big\{(s-np)^{r-1} N\mu_1(s) + \sum_{i=0}^{r-3} \binom{r-1}{i} \mu_{i+2}(s) + \tag{5.10}$$
$$+ \sum_{i=0}^{r-2} \binom{r-1}{i} \big[\big(n(p-q) - Np\big)\mu_{i+1}(s) + (N-n)npq\,\mu_i(s)\big]\Big\},$$

$$\mu_1(s) = \frac{s}{N}(Nq - n + s)p_s, \tag{5.11}$$

$$\mu_r = (N-r+1)^{-1}\Big\{\sum_{i=0}^{r-3} \binom{r-1}{i} \mu_{i+2} + \tag{5.12}$$
$$+ \sum_{i=0}^{r-2} \binom{r-1}{i} \big[\big(n(p-q) - Np\big)\mu_{i+1} + (N-n)npq\,\mu_i\big]\Big\}.$$

Die Formeln (5.2)-(5.4), (5.6)-(5.8), (5.10)-(5.12) sind schon seit längerer Zeit bekannt (siehe z.B. [4]).

Man erhält aus (5.1) die negative Binomialverteilung, wenn p durch $-p/q$, n durch $-n$ und q durch $1/q$ ersetzt wird. Dann ist

$$p_k = \mathrm{P}(X = k) = (-1)^k \binom{-n}{k} p^k q^n, \qquad q = 1-p, \; k = 0, 1, 2, \ldots, \tag{5.13}$$

und die Formeln (5.2)-(5.4) ergeben:

$$\mu_r(s) = \left(s - n\frac{p}{q}\right)^{r-1} \mu_1(s) + \frac{p}{q} \sum_{i=0}^{r-2} \binom{r-1}{i} \left(\frac{n}{q}\mu_i(s) + \mu_{i+1}(s)\right), \tag{5.14}$$

$$\mu_1(s) = \frac{s}{q} p_s, \tag{5.15}$$

$$\mu_r = \frac{p}{q} \sum_{i=0}^{r-2} \binom{r-1}{i} \left(\frac{n}{q}\mu_i + \mu_{i+1}\right). \tag{5.16}$$

Wird nun in (5.13) $p = \eta/(1+\eta)$, $n = \lambda/\eta$, $q = 1/(1+\eta)$, $p/q = \eta$ angenommen, so erhält man die sogenannte *Pólya-Eggenbergersche Verteilung*:

$$p_k = \mathrm{P}(X = k) = \frac{\Gamma(\lambda/\eta + k)}{k!\,\Gamma(\lambda/\eta)} \left(\frac{\eta}{1+\eta}\right)^k (1+\eta)^{-\lambda/\eta} \tag{5.17}$$

$$= \binom{\lambda/\eta + k - 1}{k} \left(\frac{\eta}{1+\eta}\right)^k (1+\eta)^{-\lambda/\eta}, \qquad k = 0, 1, 2, \ldots$$

Die Formeln (5.14)-(5-16) ergeben dann:

$$\mu_r(s) = (s-\lambda)^{r-1}\mu_1(s) + \eta \sum_{i=0}^{r-2} \binom{r-1}{i} \left(\frac{\lambda}{\eta}(1+\eta)\mu_i(s) + \mu_{i+1}(s)\right),$$

$$\mu_1(s) = s(1+\eta)p_s,$$

$$\mu_r = \eta \sum_{i=0}^{r-2} \binom{r-1}{i} \left(\frac{\lambda}{\eta}(1+\eta)\mu_i + \mu_{i+1}\right).$$

Die geometrische Verteilung ist ein Spezialfall der negativen Binomialverteilung (Pólya-Eggenbergerschen Verteilung), wenn $n = 1$ ($\lambda = \eta$) ist. Wir haben nämlich

$$p_k = \mathrm{P}(X = k) = qp^k, \qquad k = 0, 1, 2, \ldots, \tag{5.18}$$

oder

$$p_k = \mathrm{P}(X = k) = \frac{\eta^k}{(1+\eta)^{k+1}}, \qquad k = 0, 1, 2, \ldots, \tag{5.19}$$

und dann

$$(5.20) \quad \mu_r(s) = \left(s - \frac{p}{q}\right)^{r-1} \mu_1(s) + \frac{p}{q} \sum_{i=0}^{r-2} \binom{r-1}{i} \left(\frac{1}{q} \mu_i(s) + \mu_{i+1}(s)\right),$$

$$(5.21) \quad \mu_1(s) = \frac{s}{q} p_s,$$

$$(5.22) \qquad \mu_r = \frac{p}{q} \sum_{i=0}^{r-2} \binom{r-1}{i} \left(\frac{1}{q} \mu_i + \mu_{i+1}\right)$$

oder

$$\mu_r(s) = (s-\eta)^{r-1} \mu_1(s) + \eta \sum_{i=0}^{r-2} \binom{r-1}{i} \big((1+\eta)\mu_i(s) + \mu_{i+1}(s)\big),$$

$$\mu_1(s) = s(1+\eta) p_s,$$

$$\mu_r = \eta \sum_{i=0}^{r-2} \binom{r-1}{i} \big((1+\eta)\mu_i + \mu_{i+1}\big).$$

Wir berechnen jetzt als Beispiel die vier ersten Zentralmomente aus den Formeln (5.14)-(5.16) und (5.20)-(5.22).

Die negative Binomialverteilung:

$\mu_1(s)$ wie in (5.15),

$$\mu_2(s) = n \frac{p}{q^2} \mu_0(s) + \left(s - \frac{p}{q}(n-1)\right) \mu_1(s),$$

$$\mu_3(s) = n \frac{p}{q^2} \mu_0(s) + \left(\left(s - n\frac{p}{q}\right)^2 + \frac{p}{q^2}(q+2n)\right) \mu_1(s) + 2\frac{p}{q} \mu_2(s),$$

$$\mu_4(s) = n \frac{p}{q^2} \mu_0(s) + \left(\left(s - n\frac{p}{q}\right)^3 + \frac{p}{q^2}(q+3n)\right) \mu_1(s) + 3\frac{p}{q^2}(q+n)\mu_2(s) + 3\frac{p}{q} \mu_3(s),$$

$$\mu_2 = n \frac{p}{q^2},$$

$$\mu_3 = n \frac{p}{q^2} + 2\frac{p}{q} \mu_2 = n \frac{p}{q^3}(1+p),$$

$$\mu_4 = n \frac{p}{q^2} + 3\frac{p}{q^2}(q+n)\mu_2 + 3\frac{p}{q}\mu_3 = n\frac{p}{q^4}\big(q^2 + 3p(q+n) + 3p(1+p)\big).$$

Die geometrische Verteilung:

$\mu_1(s)$ wie in (5.21),

$$\mu_2(s) = \frac{p}{q^2}\mu_0(s) + s\mu_1(s),$$

$$\mu_3(s) = \frac{p}{q^2}\mu_0(s) + \left(\left(s-\frac{p}{q}\right)^2 + \frac{p}{q^2}(q+2)\right)\mu_1(s) + 2\frac{p}{q}\mu_2(s),$$

$$\mu_4(s) = \frac{p}{q^2}\mu_0(s) + \left(\left(s-\frac{p}{q}\right)^3 + \frac{p}{q^2}(q+3)\right)\mu_1(s) + 3\frac{p}{q^2}(q+1)\mu_2(s) + 3\frac{p}{q}\mu_3(s),$$

$$\mu_2 = \frac{p}{q^2},$$

$$\mu_3 = \frac{p}{q^2} + 2\frac{p}{q}\mu_2 = \frac{p}{q^3}(1+p),$$

$$\mu_4 = \frac{p}{q^2} + 3\frac{p}{q^2}(q+1)\mu_2 + 3\frac{p}{q}\mu_3 = \frac{p}{q^4}\big(q^2+3p(q+1)+3p(1+p)\big).$$

5.2. Wir stellen jetzt eine Methode dar, mit deren Hilfe man die gewöhnlichen Momente für die logarithmische Verteilung aus den Zentralmomenten der negativen Binomialverteilung bekommen kann. Wir bemerken nämlich, daß wegen

$$\sum_{k=1}^{\infty} \mathrm{P}(X = k) = 1-q^n,$$

wobei die Wahrscheinlichkeit $\mathrm{P}(X = k)$ durch (5.13) gegeben ist, der Ausdruck

$$\mathrm{P}(Y = l) = \frac{(-1)^l}{1-q^n}\binom{-n}{l}p^l q^n, \quad l = 1, 2, 3, \ldots, \tag{5.23}$$

die Wahrscheinlichkeit der sogenannten abgeschnittenen (gestutzten) negativen Binomialverteilung darstellt. Die Zentralmomente dieser Verteilung folgen aus der Formel (5.14), indem man (5.14) durch $1-q^n$ dividiert. Es ist bekannt, daß dann, wenn n gegen Null strebt, die Verteilung (5.23) die logarithmische Verteilung

$$\mathrm{P}(X = l) = A\,\frac{p^l}{l}, \quad l = 1, 2, 3, \ldots,$$

ergibt, wobei

$$A = -\frac{1}{\ln q}, \quad q = 1-p, \tag{5.24}$$

ist (siehe [5], S. 260, oder [8], S. 131-132). So folgt für die abgeschnittene negative Binomialverteilung

$$\mu_r(s) = \left(s - n\frac{p}{q}\right)^{r-1} \frac{\mu_1(s)}{1-q^n} + \frac{p}{q(1-q^n)} \sum_{i=0}^{r-2} \binom{r-1}{i} \left(\frac{n}{q}\mu_i(s) + \mu_{i+1}(s)\right)$$

und dann die Formel für die unvollständigen gewöhnlichen Momente $a_r(s)$ der logarithmischen Verteilung

$$\text{(5.25)} \qquad a_r(s) = s^{r-1}a_1(s) + \frac{p}{q}\sum_{i=0}^{r-2}\binom{r-1}{i}a_{i+1}(s)$$

$$= \frac{p}{q}\left(As^{r-1}p^{s-1} + \sum_{i=0}^{r-2}\binom{r-1}{i}a_{i+1}(s)\right),$$

weil wegen der Bedingung $n \to 0$ es gilt

$$\frac{\mu_r(s)}{1-q^n} = \sum_{k=s}^{\infty}\left(k - n\frac{p}{q}\right)^r \frac{p_k}{1-q^n} \to a_r(s)$$

und aus der Definition $a_1(s) = Ap^s/q$ ist. (Siehe auch Gerstenkorn [4], S. 22-24 und S. 33, wo die anderen Methoden für die unvollständigen Momente, und Kendall und Stuart [8], S. 133, wo die Formeln für die vier ersten vollständigen Momente angegeben sind.)

Setzt man $s = 1$ voraus, so erhält man die Formel für die vollständigen gewöhnlichen Momente a_r der logarithmischen Verteilung

$$a_r = \frac{p}{q}\left(A + \sum_{i=0}^{r-2}\binom{r-1}{i}a_{i+1}\right),$$

wobei A in (5.24) bereits definiert wurde.

Zur Anschauung werden hier die ersten vier unvollständigen Momente aus der Formel (5.25) berechnet (das erste Moment direkt aus der Definition):

$$a_1(s) = \sum_{l=s}^{\infty} lp_l = A\sum_{l=s}^{\infty}p^l = A\left(\frac{p}{q} - \sum_{l=1}^{s-1}p^l\right)$$

$$= A\left(\frac{p}{q} - \frac{p}{q}(1-p^{s-1})\right) = A\frac{p^s}{q},$$

$$a_2(s) = \frac{p}{q}\left(Asp^{s-1} + a_1(s)\right) = A\frac{p^s}{q}\left(s + \frac{p}{q}\right),$$

$$a_3(s) = \frac{p}{q}\left(As^2p^{s-1}+a_1(s)+2a_2(s)\right) = A\frac{p^s}{q}\left(s^2+\frac{p}{q}+2\frac{p}{q}\left(s+\frac{p}{q}\right)\right),$$

$$\begin{aligned} a_4(s) &= \frac{p}{q}\left(As^3p^{s-1}+a_1(s)+3a_2(s)+3a_3(s)\right) \\ &= A\frac{p^s}{q}\left\{s^3+\frac{p}{q}+3\frac{p}{q}\left(s+\frac{p}{q}\right)+3\frac{p}{q}\left(s^2+\frac{p}{q}+2\frac{p}{q}\left(s+\frac{p}{q}\right)\right)\right\}. \end{aligned}$$

Die anderen Formeln für die unvollständigen und vollständigen gewöhnlichen Momente der logarithmischen Verteilung waren schon eher bekannt (siehe [4], S. 33, (2.45), (2.47)).

6. Anwendungen. Die erlangten Ergebnisse werden nun zur Ermittlung der absoluten Zentralmomente der Pólya-Verteilung (und der Verteilungen, die man aus dieser erhalten kann) angewendet. Im Spezialfall erhält man auch die Formeln für die mittlere Abweichung der entsprechenden Verteilungen.

Es bezeichne β_r das absolute Zentralmoment der zufälligen Veränderlichen X mit dem Erwartungswert $E(X) = a$, d.h. es gilt $\beta_r = E(|X-a|^r)$.

Es sei bemerkt, daß im Fall der diskreten Veränderlichen mit der Wahrscheinlichkeitsfunktion $p_k = P(X=k)$ für β_r (ungerades r) folgt

$$(6.1)\qquad \begin{aligned} \beta_r &= \sum_{k=-\infty}^{\infty}|k-a|^r p_k = -\sum_{k=-\infty}^{l-1}(k-a)^r p_k + \sum_{k=l}^{\infty}(k-a)^r p_k \\ &= 2\sum_{k=l}^{\infty}(k-a)^r p_k - \mu_r = 2\mu_r(l)-\mu_r, \end{aligned}$$

wobei $l = [a]+1$ ist und $[a]$ den Gesamtteil der Zahl a bezeichnet.

Im Spezialfall, wenn $r = 1$ ist, erhält man die Formel für die mittlere Abweichung β_1:

$$(6.2)\qquad \beta_1 = 2\mu_1(l).$$

Wenn man jetzt die Formeln (4.10), (4.11) und $a = np$ in Zusammenhang mit (6.1) sieht, dann erhält man die Formel für die absoluten Zentralmomente der Pólya-Verteilung:

$$(6.3)\qquad \begin{aligned} \beta_r &= \left(1+(r-1)a\right)^{-1}\Big\{2(l-np)^{r-1}\mu_1(l) - \\ &-a\sum_{i=0}^{r-3}\binom{r-1}{i}\left(2\mu_{i+2}(l)-\mu_{i+2}\right) + \\ &+\sum_{i=0}^{r-2}\binom{r-1}{i}\Big(\left(na(q-p)-p\right)\left(2\mu_{i+1}(l)-\mu_{i+1}\right)+npq(1+na)\left(2\mu_i(l)-\mu_i\right)\Big)\Big\}. \end{aligned}$$

Wenn wir in (6.3)

$$2\mu_{i+k}(l)-\mu_{i+k} = m_{i+k} \qquad (k = 0, 1, 2)$$

setzen, dann gilt nach Berücksichtigung von (6.2) die Formel

$$(6.4)\qquad \beta_r = \big(1+(r-1)a\big)^{-1}\Big\{(l-np)^{r-1}\beta_1 - a\sum_{i=0}^{r-3}\binom{r-1}{i} m_{i+2} + \\ + \sum_{i=0}^{r-2}\binom{r-1}{i}\Big(\big(na(q-p)-p\big)m_{i+1}+npq(1+na)m_i\Big)\Big\},$$

wobei

$$(6.5)\qquad \beta_1 = 2l\big(q+(n-l)a\big)p_l.$$

ist.

Für die Gestalt (4.13) der Pólya-Verteilung gilt die folgende Abänderung der Formeln (6.4) und (6.5):

$$(6.6)\qquad \beta_r = \big(N+(r-1)S\big)^{-1}\Big\{(l-np)^{r-1}N\beta_1 - S\sum_{i=0}^{r-3}\binom{r-1}{i} m_{i+2} + \\ + \sum_{i=0}^{r-2}\binom{r-1}{i}\Big(\big(nS(q-p)-Np\big)m_{i+1}+npq(N+nS)m_i\Big)\Big\},$$

wobei

$$(6.7)\qquad \beta_1 = \frac{2}{N}\,l\big(qN+(n-l)S\big)p_l$$

ist.

Die Formeln (6.4) und (6.5) können nun als Ausgangsformeln für viele andere dienen. Nämlich, wenn $a = 0$ ist, so erhält man die Formel für die absoluten Zentralmomente der Binomialverteilung (5.1),

$$\beta_r = (l-np)^{r-1}\beta_1 + p\sum_{i=0}^{r-2}\binom{r-1}{i}(nqm_i - m_{i+1}),$$

wobei

$$(6.8)\qquad \beta_1 = 2lqp_l, \qquad l = [np]+1,$$

ist. Die Formel (6.8) ist schon seit dem Jahr 1957 bekannt und wurde in [6] angegeben.

Wenn in (6.7)

$$\lim_{n\to\infty} np = \lambda$$

vorausgesetzt wird, dann erhält man die Formel für die absoluten Zentralmomente der Poissonschen Verteilung (5.5),

$$\beta_r = (l-\lambda)^{r-1}\beta_1 + \lambda \sum_{i=0}^{r-2} \binom{r-1}{i} m_i,$$

wobei

(6.9) $$\beta_1 = 2lp_l, \quad l = [\lambda]+1,$$

ist.

Wenn man nun in (6.6) den Wert $S = -1$ annimmt, so erhält man die Formel für die absoluten zentralen Momente der hypergeometrischen Verteilung (5.9),

$$\beta_r = \big(N-(r-1)\big)^{-1}\Big\{(l-np)^{r-1}N\beta_1 + \sum_{i=0}^{r-3}\binom{r-1}{i} m_{i+2} +$$

$$+ \sum_{i=0}^{r-2}\binom{r-1}{i}\Big(npq(N-n)m_i - \big(n(q-p)+Np\big)m_{i+1}\Big)\Big\},$$

wobei

(6.10) $$\beta_1 = \frac{2}{N} l(Nq-n+l)p_l, \quad l = [np]+1,$$

ist.

Die obenerwähnte Methode liefert uns die Formeln für die negative Binomialverteilung (5.13),

$$\beta_r = \left(l - n\frac{p}{q}\right)^{r-1}\beta_1 + \frac{p}{q}\sum_{i=0}^{r-2}\binom{r-1}{i}\left(m_{i+1} + \frac{n}{q} m_i\right),$$

wobei

(6.11) $$\beta_1 = \frac{2}{q} lp_l, \quad l = \left[n\frac{p}{q}\right]+1,$$

ist. In analoger Weise erhält man für die Pólya-Eggenberger-Verteilung (5.17),

$$\beta_r = (l-\lambda)^{r-1}\beta_1 + \eta\sum_{i=0}^{r-2}\binom{r-1}{i}\left(m_{i+1} + \frac{\lambda}{\eta}(1+\eta)m_i\right),$$

wobei $\beta_1 = 2(1+\eta)lp_l$, $l = [\lambda]+1$, ist und dann, für die geometrische Verteilung (5.18),

$$\beta_r = \left(l - \frac{p}{q}\right)^{r-1}\beta_1 + \frac{p}{q}\sum_{i=0}^{r-2}\binom{r-1}{i}\left(m_{i+1} + \frac{1}{q} m_i\right),$$

wobei

$$\beta_1 = \frac{2}{q} l p_l, \quad l = \left[\frac{p}{q}\right] + 1, \tag{6.12}$$

ist. Für den Fall (5.19) ist

$$\beta_r = (l-\eta)^{r-1} \beta_1 + \eta \sum_{i=0}^{r-2} \binom{r-1}{i} \left(m_{i+1} + (1+\eta) m_i\right)$$

mit $\beta_1 = 2(1+\eta) l p_l$, $l = [\eta] + 1$.

Die Bedeutung der hier gegebenen Formeln und der Methode sie zu ermitteln, ist leicht zu ermessen, wenn man einen Vergleich dieser Methode mit der Methode der Berechnung, die direkt aus der Definitionsformel erfolgt, beispielsweise der mittleren Abweichung der Binomialverteilung, anstellt, selbst wenn die Werte der Parameter sorgfältig angepasst sind. Man findet diese Berechnung in [5], S. 244 (siehe auch die Bemerkungen in [12], S. 92 und 94).

Die Formeln für die mittlere Abweichung der obenerwähnten Verteilungen sind schon seit einiger Zeit bekannt. Am frühesten ist die Formel (6.8) für die mittlere Abweichung der binomialen Verteilung bekannt geworden. Sie wurde von Uspensky [13] ohne Berücksichtigung der Formel (6.2), dann von Frame [2], Koźniewska [9], Johnson [6] und Kamat [7] gegeben. Koźniewska [9] hat die Formel (6.5) für die mittlere Abweichung der Pólya-Verteilung gefunden (Spezialfälle: die Formeln (6.8) und (6.10) für die binomiale und hypergeometrische Verteilung). Die mittlere Abweichung (6.9) für die Poissonsche Verteilung wurde von Crow [1] und Kamat [7] betrachtet. Die Formel (6.10) für die mittlere Abweichung der hypergeometrischen Verteilung ist außer der Arbeit von Koźniewska [9] auch in Ramasubban [11] und Kamat [7] zu finden. In den Arbeiten der letzten Verfasser sind auch die Formeln (6.11) für die negativ-binomiale Verteilung und (6.12) für die geometrische Verteilung publiziert worden.

Wir möchten noch bemerken, daß ein Zusammenhang zwischen der mittleren Abweichung der Pólya-Verteilung und der Varianz μ_2^B der binomialen Verteilung und der Varianz μ_2^P der Pólya-Verteilung festzustellen ist. Wenn wir nämlich in die Formel (6.5) die Wahrscheinlichkeit (4.2) und $l-1 = [np]$ einsetzen, bekommen wir

$$\beta_1 = 2(n-[np])(p+[np]a) p_{l-1}. \tag{6.13}$$

Wenn jetzt $[np] = np$ ist, dann folgt daraus

$$\beta_1 = 2\mu_2^B (1+na) p_{l-1} \tag{6.14}$$

oder

$$\beta_1 = 2\mu_2^P (1+a) p_{l-1}. \tag{6.15}$$

Im Gegenfall, d.h. wenn $[np] \neq np$ ist, können die Formeln (6.14) und (6.15) als Approximationsformeln dienen. Aus den Formeln (6.13)-(6.15) sind auch analogische Formeln für die binomiale, Poissonsche, hypergeometrische, negativ-binomiale und geometrische Verteilung zu entwickeln. Sie sind von Kamat [7] mit einem besonderen Verfahren erlangt worden.

Noch früher hat Frame [2] auf die Beziehung $2\mu_2^B p_{l-1}$ in der Binomialverteilung und Crow [1] auf die Beziehung $2\lambda p_{l-1}$ in der Poissonschen Verteilung hingewiesen.

Literaturverzeichnis

[1] E. L. Crow, *The mean deviation of the Poisson distribution*, Biometrika 45 (1958), S. 556-559.

[2] J. S. Frame, *Mean deviation of the binomial distribution*, Amer. Math. Monthly 52 (1945), S. 377-379.

[3] R. Frisch, *Recurrence formulae for the moments of the point binomial distribution*, Biometrika 17 (1925), S. 165-171.

[4] T. Gerstenkorn, *The recurrence relations for the moments of the discrete probability distributions*, Dissertationes Mathematicae 83 (1971).

[5] – und T. Śródka, *Kombinatoryka i rachunek prawdopodobieństwa*, PWN, Warszawa 1972.

[6] N. L. Johnson, *A note on the mean deviation of the binomial distribution*, Biometrika 44 (1957), S. 532-533.

[7] A. R. Kamat, *Incomplete and absolute moments of some discrete distributions, Classical and contagious discrete distributions*, Proceedings of the International Symposium, McGill University, Montreal, Canada, August 15th - 20th 1963, edited by G. P. Patil, Statistical Publishing Society, Calcutta 1965, S. 45-64.

[8] M. G. Kendall and A. Stuart, *The advanced theory of statistics*. Vol. 1, *Distribution theory*, Griffin, London 1958.

[9] I. Koźniewska, *Pierwszy absolutny moment centralny dla rozkladu Pólyi* (*The first absolute central moment for Pólya's distribution*), Zastosow. Matem. 1 (1954), S. 206-211.

[10] G. Mühlbach, *Rekursionsformeln für die zentralen Momente der Pólya- und der Beta-Verteilung*, Metrika 19 (1972), S. 171-177.

[11] T. A. Ramasubban, *The mean difference and the mean deviation of some discontinuous distributions*, Biometrika 45 (1958), S. 549-556.

[12] A. Rényi, *Wahrscheinlichkeitsrechnung mit einem Anhang über Informationstheorie*, VEB Deutscher Verlag der Wissenschaften, Berlin 1962.

[13] J. V. Uspensky, *Introduction to mathematical probability*, McGraw Hill, New York 1937, S. 176-177.

MATHEMATISCHES INSTITUT
DER UNIVERSITÄT ŁÓDŹ

Eingegangen am 15. 12. 1973

T. GERSTENKORN (ŁÓDŹ)

UWAGI O MOMENTACH CENTRALNYCH NIEKOMPLETNYCH I ABSOLUTNYCH ROZKŁADU PÓLYI

STRESZCZENIE

W pracy podane są wzory rekurencyjne na momenty centralne niekompletne i kompletne rozkładu Pólyi, uzyskane na podstawie twierdzenia A. R. Kamata. Dla przykładu obliczono za ich pomocą cztery pierwsze momenty. Z tych zasadniczych wzorów otrzymuje się — jako przypadki szczególne — wzory na omawiane momenty dla rozkładu dwumianowego, Poissona, hipergeometrycznego, ujemnego dwumianowego, Pólyi-Eggenbergera, geometrycznego oraz wzory rekurencyjne na niekompletne i kompletne momenty zwykłe rozkładu logarytmicznego.

Momenty niekompletne wykorzystuje się do wyprowadzenia prostych wzorów na momenty centralne absolutne dowolnego rzędu (w szczególności odchylenia przeciętnego) omawianych rozkładów.