Werk

Titel: Vergleich der hypergeometrischen mit der Binomial-Verteilung

Autor: Uhlmann, W.

Jahr: 1966

PURL: https://resolver.sub.uni-goettingen.de/purl?358794056_0010|log23

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

Vergleich der hypergeometrischen mit der Binomial-Verteilung

Von W. UHLMANN, Würzburg[1])

Zusammenfassung: Es sei L die Operations-Charakteristik beim Testen einer Hypothese über eine unbekannte Wahrscheinlichkeit p eines Ereignisses A. L ist also die Wahrscheinlichkeit dafür, daß höchstens c Elemente mit A in der Stichprobe auftreten. Wird die Stichprobe ohne Zurücklegen gezogen, so hängt L vom Umfang N der Grundgesamtheit ab. Es wird untersucht, für welche p die Operations-Charakteristik L monoton von N abhängt. Den Abschluß bilden Vergleiche der Operations-Charakteristik bei der Binomial-Verteilung mit L und deren Annäherung durch die Poisson-Verteilung.

Summary: Let L be the operating characteristic for testing a hypothesis concerning an unknown probability p of an event A. Then L is the probability of being no more than c elements with A in the sample. When the sample is drawn without replacement L depends on the size N of the population. Investigation is made for which p the operating characteristic L depends monotone on N. Finally the operating characteristic of the binomial distribution is compared with L and with the approximation by the Poisson distribution.

1. Einleitung

Es sei $W(A) = p$ die Wahrscheinlichkeit dafür, daß ein aus einer Grundgesamtheit zufällig gezogenes Element eine gewisse Eigenschaft A hat. Um die Nullhypothese $p \leqslant \pi_0$ gegen die Alternative $p > \pi_0$ zu testen, benutzt man die Anzahl x der Elemente mit der Eigenschaft A in einer Stichprobe vom Umfang $n \geqslant 1$. Die Größe der vorgegebenen Zahl π_0 ist für die folgenden Untersuchungen nicht von Belang. Die Nullhypothese wird genau dann angenommen, wenn $x \leqslant c$ ausfällt, wobei die Annahmezahl c eine natürliche Zahl mit $0 \leqslant c \leqslant n - 1$ sei.

Die statistische Qualitätskontrolle ist eine besonders wichtige Anwendungsmöglichkeit für diesen Test. Man hat dort über die Annahme oder Ablehnung einer Partie von Waren auf Grund einer Stichprobe zu ent-

[1]) Prof. Dr. WERNER UHLMANN, Institut für Statistik der Universität Würzburg.

scheiden, wobei bei jedem Stichprobenelement nur festgestellt wird, ob es „schlecht" im Sinne der Lieferbedingungen ist oder nicht. Wir wollen daher die Eigenschaften dieses Testes nicht durch die Gütefunktion, sondern durch die Operations-Charakteristik L ausdrücken, also durch die Wahrscheinlichkeit $W_p(x \leqslant c)$ in Abhängigkeit von p für die Annahme der Nullhypothese.

Wird die Stichprobe „mit Zurücklegen" gezogen oder ist die Grundgesamtheit unendlich groß, so ist die zufällige Variable x nach der Binomial-Verteilung $\mathrm{Bi}(n, p)$ verteilt und die Operations-Charakteristik lautet

$$L_{n,c}(p) = W_p(x \leqslant c) = \sum_{m=0}^{c} \binom{n}{m} p^m (1-p)^{n-m}. \tag{1}$$

Beim Ziehen „ohne Zurücklegen" dagegen ist x nach der hypergeometrischen Verteilung $H(N, n, p)$ verteilt und die Operations-Charakteristik

$$L_{N,n,c}(p) = W_p(x \leqslant c) = \tag{2}$$

$$= \begin{cases} 1 & \text{für} \quad 0 \leqslant p \leqslant \dfrac{c}{N} \\ \displaystyle\sum_{m=0}^{c} \binom{pN}{m}\binom{N-pN}{n-m} \Big/ \binom{N}{n} & \text{für} \quad \dfrac{c}{N} \leqslant p \leqslant 1 - \dfrac{n-c-1}{N} \\ 0 & \text{für} \quad 1 - \dfrac{n-c-1}{N} \leqslant p \leqslant 1 \end{cases}$$

hängt noch von einem weiteren Parameter, der Anzahl N der Elemente der Grundgesamtheit (im Beispiel: Umfang der Partie) ab. Da pN die Anzahl der Elemente mit der Eigenschaft A in der Grundgesamtheit ist, ist $L_{N,n,c}(p)$ zunächst nur für solche p definiert, für die pN ganzzahlig ist. Um zu einfacher formulierbaren Sätzen zu gelangen, sehen wir aber $L_{N,n,c}(p)$ durch (2) als für alle reellen p zwischen 0 und 1 definiert an.

Bei dem obigen Test kann man vier reelle Zahlen α, p_α, β, p_β mit $1 > \alpha > 1/2 > \beta > 0$ und $0 < p_\alpha < p_\beta < 1$ vorgeben und den Stichprobenumfang n und die Annahmezahl c so bestimmen, daß n möglichst klein und

$$L_{n,c}(p_\alpha) \geqslant \alpha \qquad \text{und} \qquad L_{n,c}(p_\beta) \leqslant \beta \tag{3}$$

ist. Wie zum Beispiel in [3] ausführlich dargestellt ist, kann man sich zur Berechnung von n und c der Approximation der Binomial-Verteilung durch die Poisson-Verteilung oder durch die Normalverteilung bedienen und die geforderten Ungleichungen (3) mit Hilfe des Zusammenhanges zwischen der Binomial-Verteilung und der F-Verteilung nachprüfen.

Nun werden in der Praxis Stichproben stets „ohne Zurücklegen" gezogen. Wenn der Stichprobenumfang n erheblich kleiner als N ist, kann man aus einer z. B. bei FELLER [1], Seite 57 wiedergegebenen Abschätzung aus der Gültigkeit von (3) folgern, daß für diese n und c auch $L_{N,n,c}(p)$ der (3) entsprechenden Bedingung wenigstens näherungsweise genügt. Es fragt sich aber, ob dieser Schluß auch für den Fall noch richtig ist, daß N nur wenig größer als n ist. Da man beim „Ziehen ohne Zurücklegen" etwas genauere Informationen über die Grundgesamtheit erhält als beim „Ziehen mit Zurücklegen" (siehe z. B. die etwas kleinere Streuung von $H(N, n, p)$ gegenüber der bei $\mathrm{Bi}(n, p)$), ist von vornherein zu vermuten, daß $L_{N,n,c}(p)$ die Bedingungen (3) erst recht erfüllt, doch fehlen m. E. bisher in der Literatur die Angabe der Voraussetzungen ebenso wie die zugehörigen Beweise. Lediglich der Fall $c = 0$, und damit natürlich auch der Fall $c = n - 1$, ist in allen einschlägigen Lehrbüchern wiedergegeben:

$$\left.\begin{aligned} L_{N,n,0}(p) &< L_{n,0}(p) \\ L_{N,n,n-1}(p) &> L_{n,n-1}(p) \end{aligned}\right\} \quad \text{für} \quad 0 < p < 1. \tag{4}$$

Es soll daher im folgenden (siehe Satz 5) gezeigt werden, daß $L_{N,n,c}(p)$ die Kurve $L_{n,c}(p)$ nur in nächster Nähe des 50%-Punktes von $L_{n,c}(p)$ schneidet, der ungefähr gleich $c/(n-1)$ ist, wie man etwa der Approximation durch die Poisson-Verteilung oder durch die Normalverteilung entnehmen kann (siehe auch Satz 6). Um den Beweis zu führen, wird nachgewiesen (siehe Satz 4), daß die bekannte Konvergenz von $L_{N,n,c}(p)$ gegen $L_{n,c}(p)$ für $N \to \infty$ außerhalb einer kleinen Umgebung von $c/(n-1)$ monoton ist.

2. Monotonie der Operations-Charakteristik

Um ohne großen Rechenaufwand Aussagen über die Operations-Charakteristik $L_{N,n,c}(p)$ machen zu können, nutzen wir aus, daß es sich im entscheidenden Intervall $[c/N, 1 - (n - c - 1)/N]$ um ein Polynom in p vom Grad n handelt. Wir betrachten daher die Hilfsfunktion

$$\tilde{L}_{N,n,c}(p) = \sum_{m=0}^{c} \binom{pN}{m} \binom{N - pN}{n - m} \Big/ \binom{N}{n}, \tag{5}$$

die im obigen Intervall mit $L_{N,n,c}(p)$ übereinstimmt, für alle reellen p und die natürlichen Zahlen N, n und c mit $0 \leqq c < n \leqq N$.

Es ist

$$\tilde{L}_{N,n,c}\left(\frac{i}{N}\right) = \begin{cases} 1 & \text{für} \quad i = 0, 1, \ldots, c \\ 0 & \text{für} \quad i = N - (n - c - 1), \ldots, N. \end{cases}$$

Daher hat die erste Ableitung $(d/dp)\,\tilde{L}_{N,n,c}(p)$ als Polynom in p vom Grad $(n-1)$ genau die $(n-1)$ verschiedenen Nullstellen $p'_1, p'_2, \ldots, p'_{n-1}$ mit

$$0 < p_1' < \frac{1}{N} < p_2' < \frac{2}{N} < \ldots < \frac{c-1}{N} < p_c' < \frac{c}{N}$$

und

$$\frac{N-n+c+1}{N} < p_{c+1}' < \frac{N-n+c+2}{N} < \ldots < \frac{N-1}{N} < p_{n-1}' < 1.$$

Daraus folgt insbesondere

Satz 1: Die Operations-Charakteristik $L_{N,n,c}(p)$ ist im Intervall $(c/N, 1-(n-c-1)/N)$ streng monoton abnehmend.

Es sei bemerkt, daß dieser für die sinnvolle Anwendungsmöglichkeit des Testes unabdingbare Satz für den Fall, daß $p\,N$ ganzzahlig ist, üblicherweise durch Betrachtung der 1. Differenz mit der Schrittweite $1/N$ bewiesen wird.

3. Ein Hilfssatz

Als Vorbereitung für den folgenden Abschnitt wollen wir hier eine Abschätzung für die eben erwähnte 1. Differenz von $L_{N,n,c}(p)$ herleiten.

Es ist klar, daß die p_i' abwechselnd Stellen relativer Maxima und Minima sind und daher zwischen p_k' und p_{k+1}' jeweils genau eine Nullstelle p_k'' von $(d^2/dp^2)\,L_{N,n,c}(p)$ liegt $(k = 1, 2, \ldots, n-2)$.

Für $0 < c < n-1$ brauchen wir noch das Vorzeichen dieser zweiten Ableitung im Intervall (p_c', p_{c+1}'). Da $L_{N,n,c}(p)$ ersichtlich an der Stelle p_c' ein relatives Maximum und an der Stelle p_{c+1}' ein relatives Minimum hat, gilt für $0 < c < n-1$:

$$\frac{d^2}{dp^2} L_{N,n,c}(p) \begin{cases} < 0 & \text{für} \quad p_c'' \leqslant p < p_c'' \\ > 0 & \text{für} \quad p_c'' < p \leqslant p_{c+1}'. \end{cases} \tag{6}$$

Für die 1. Differenz von $L_{N,n,c}(p)$ beweisen wir für alle reellen p durch vollständige Induktion nach c:

$$L_{N,n,c}(p) - L_{N,n,c}\left(p + \frac{1}{N}\right) = \binom{p\,N}{c}\binom{N-p\,N-1}{n-c-1} \Big/ \binom{N}{n}. \tag{7}$$

Offenbar stimmt die Behauptung für $c = 0$, und der Schluß von $(c-1)$ auf c mit $1 \leqslant c \leqslant n-1$ ergibt sich aus der elementaren Rechnung:

$$\binom{p\,N}{c-1}\binom{N-p\,N-1}{n-c} + \binom{p\,N}{c}\binom{N-p\,N}{n-c} - \binom{p\,N+1}{c}\binom{N-p\,N-1}{n-c} =$$

$$= \binom{p\,N}{c}\binom{N-p\,N}{n-c} - \binom{p\,N}{c}\binom{N-p\,N-1}{n-c} = \binom{p\,N}{c}\binom{N-p\,N-1}{n-c-1}.$$

Für $1 < c < n - 1$ ergibt sich aus (7) sofort die 2. Differenz:

$$\tilde{L}_{N,n,c}\left(p - \frac{1}{N}\right) - 2\,\tilde{L}_{N,n,c}(p) + \tilde{L}_{N,n,c}\left(p + \frac{1}{N}\right) =$$

$$= \frac{\binom{pN-1}{c-1}\binom{N-pN-1}{n-c-2}}{c(n-c-1)\binom{N}{n}} N\{(n-1)p - c\}. \tag{8}$$

Die 2. Differenz verschwindet also für $p = c/(n-1)$ und daher liegen die drei Punkte

$$\left(\frac{c}{n-1} - \frac{1}{N}, \tilde{L}_{N,n,c}\left(\frac{c}{n-1} - \frac{1}{N}\right)\right), \quad \left(\frac{c}{n-1}, \tilde{L}_{N,n,c}\left(\frac{c}{n-1}\right)\right),$$

$$\left(\frac{c}{n-1} + \frac{1}{N}, \tilde{L}_{N,n,c}\left(\frac{c}{n-1} + \frac{1}{N}\right)\right)$$

auf einer Geraden. Da

$$\frac{d^2}{dp^2}\,\tilde{L}_{N,n,c}(p)$$

im Intervall $[p'_c, p'_{c+1}]$ nicht mehr als eine Nullstelle hat, nämlich p''_c mit

$$\frac{c}{n-1} - \frac{1}{N} < p''_c < \frac{c}{n-1} + \frac{1}{N},$$

folgt daraus zusammen mit (6), daß die Sehnen durch die Punkte

$$\left(p - \frac{1}{N}, \tilde{L}_{N,n,c}\left(p - \frac{1}{N}\right)\right) \quad \text{und} \quad (p, \tilde{L}_{N,n,c}(p))$$

für alle p mit

$$\frac{c}{N} \leqslant p \leqslant \frac{c}{n-1}$$

ganz unterhalb von $\tilde{L}$ verlaufen und daß sie für

$$\frac{c}{n-1} + \frac{1}{N} \leqslant p \leqslant 1 - \frac{n-c-2}{N}$$

ganz oberhalb von $\tilde{L}$ liegen. Dieser Sachverhalt läßt sich auch so ausdrücken: Für jede reelle Zahl ε mit $0 < \varepsilon < 1$ und für jedes p mit

$$\frac{c}{N} \leqq p \leqq \frac{c}{n-1}$$

ist

$$\varepsilon\left\{\bar{L}_{N,n,c}\left(p-\frac{1}{N}\right)-\bar{L}_{N\,n,c}(p)\right\}<\bar{L}_{N,n,c}\left(p-\frac{\varepsilon}{N}\right)-\bar{L}_{N,n,c}(p). \tag{9}$$

Dagegen ist für

$$\frac{c}{n-1}+\frac{1}{N} \leqq p \leqq 1-\frac{n-c-2}{N}$$

$$\varepsilon\left\{\bar{L}_{N,n,c}\left(p-\frac{1}{N}\right)-\bar{L}_{N,n,c}(p)\right\}>\bar{L}_{N,n,c}\left(p-\frac{\varepsilon}{N}\right)-\bar{L}_{N,n,c}(p). \tag{10}$$

Setzen wir in (9) und (10) die Gleichung (7) für die 1. Differenz ein, so folgt:

Hilfssatz 1: Es sei $0 < c < n-1 < N$ *und* ε *eine reelle Zahl mit* $0 < \varepsilon < 1$. *Dann ist*

$$\bar{L}_{N,n,c}\left(p-\frac{\varepsilon}{N}\right)-\bar{L}_{N,n,c}(p)>\varepsilon\binom{pN-1}{c}\binom{N-pN}{n-c-1}\Big/\binom{N}{n} \tag{11}$$

falls

$$\frac{c}{N} \leqq p \leqq \frac{c}{n-1}$$

ist. Dagegen ist

$$\bar{L}_{N,n,c}\left(p-\frac{\varepsilon}{N}\right)-\bar{L}_{N,n,c}(p)<\varepsilon\binom{pN-1}{c}\binom{N-pN}{n-c-1}\Big/\binom{N}{n} \tag{12}$$

für

$$\frac{c}{n-1}+\frac{1}{N} \leqq p \leqq 1-\frac{n-c-2}{N}.$$

4. Monotonie der Konvergenz

Um Aussagen über die Monotonie der Konvergenz von $L_{N,n,c}(p)$ gegen $L_{n,c}(p)$ für $N \to \infty$ machen zu können, wollen wir zunächst die Schnittpunkte von $\bar{L}_{N,n,c}(p)$ mit $\bar{L}_{N+1,n,c}(p)$ betrachten, und zwar für $0 \leqq c < n \leqq N$.

Ersichtlich sind $p = 0$ und $p = 1$ die Abszissen zweier solcher Schnittpunkte. Für $c > 1$ ist $\bar{L}_{N,n,c}(p)$ an den Stellen

$$\frac{c}{N+1}, \frac{c-1}{N+1}, \frac{c-2}{N+1}, \ldots, \frac{1}{N+1}$$

immer abwechselnd > 1 beziehungsweise < 1, dagegen ist $\tilde{L}_{N+1,n,c}(p)$ dort immer $= 1$. Daher liegen in den $(c-1)$ offenen Intervallen

$$\left(\frac{1}{N+1}, \frac{2}{N+1}\right), \ldots, \left(\frac{c-1}{N+1}, \frac{c}{N+1}\right)$$

mindestens $(c-1)$ Schnittpunkt-Abszissen. Entsprechend können wir für $c < n-2$ feststellen, daß $\tilde{L}_{N,n,c}(p)$ an den Stellen

$$1 - \frac{n-c-1}{N+1}, 1 - \frac{n-c-2}{N+1}, \ldots, 1 - \frac{1}{N+1}$$

immer abwechselnd < 0 beziehungsweise > 0 ist, während

$$\tilde{L}_{N+1,n,c}\left(1 - \frac{i}{N+1}\right) = 0$$

für $i = 1, \ldots, n-c-1$ ist. Daher enthält jedes der $(n-c-2)$ offenen Intervalle

$$\left(1 - \frac{n-c-1}{N+1}, 1 - \frac{n-c-2}{N+1}\right), \ldots, \left(1 - \frac{2}{N+1}, 1 - \frac{1}{N+1}\right)$$

ebenfalls mindestens eine Schnittpunkt-Abszisse. Für $0 < c < n-1$ ist schließlich noch

$$\tilde{L}_{N+1,n,c}\left(\frac{c}{N}\right) < \tilde{L}_{N,n,c}\left(\frac{c}{N}\right) = 1$$

und

$$\tilde{L}_{N+1,n,c}\left(1 - \frac{n-c-1}{N}\right) > \tilde{L}_{N,n,c}\left(1 - \frac{n-c-1}{N}\right) = 0,$$

und es muß im Intervall

$$\left(\frac{c}{N}, 1 - \frac{n-c-1}{N}\right)$$

noch eine weitere Schnittpunkt-Abszisse liegen. Damit haben wir insgesamt mindestens n Schnittpunkte (auch für die Fälle $c = 0$, $c = 1$,

$c = n - 2$ und $c = n - 1$) ermittelt, und da $\tilde{L}_{N+1,n,c}(p) - \tilde{L}_{N,n,c}(p)$ ein Polynom höchstens vom Grad n ist, kann es keine weiteren Schnittpunkte geben.

Für $c = 0$ und $c = n - 1$ können wir daraus schon die Monotonie von $L_{N,n,c}(p)$ in Abhängigkeit von N folgern:

Satz 2: Für $c = 0 < n \leqslant N$ *ist*

$$L_{N,n,0}(p) - L_{N+1,n,0}(p) \begin{cases} = 0 & \text{für} \quad p = 0 \\ < 0 & \text{für} \quad 0 < p < 1 - \dfrac{n-1}{N+1} \\ = 0 & \text{für} \quad 1 - \dfrac{n-1}{N+1} \leqslant p \leqslant 1 \end{cases}$$

und für $1 \leqslant c = n - 1 < N$ *ist*

$$L_{N,n,n-1}(p) - L_{N+1,n,n-1}(p) \begin{cases} = 0 & \text{für} \quad 0 \leqslant p \leqslant \dfrac{n-1}{N+1} \\ > 0 & \text{für} \quad \dfrac{n-1}{N+1} < p < 1 \\ = 0 & \text{für} \quad p = 1. \end{cases}$$

Für $0 < c < n - 1$ dagegen können wir im Augenblick nur folgern:

Hilfssatz 2: Für $0 < c < n - 1 < N$ *gibt es genau eine reelle Zahl* p_0 *mit*

$$\frac{c}{N} < p_0 < 1 - \frac{n-c-1}{N}$$

und $L_{N,n,c}(p_0) = L_{N+1,n,c}(p_0)$.

Für den Fall, daß n eine ungerade Zahl > 1 und $c = (n-1)/2$ ist, läßt sich dieser Schnittpunkt p_0 explizit angeben: Es ist $p_0 = 1/2$, da ersichtlich für $c = (n-1)/2$, $p = 1/2$ und jede natürliche Zahl $N \geqslant n$

$$\sum_{m=0}^{c} \binom{pN}{m} \binom{N-pN}{n-m} = \sum_{m=c+1}^{n} \binom{pN}{m} \binom{N-pN}{n-m} = \frac{1}{2} \binom{N}{n}$$

ist. Daraus folgt zusammen mit der Definition (2) und Satz 1:

Satz 3: Wenn n *eine ungerade Zahl* > 1 *ist, so ist für* $c = (n-1)/2$

$$L_{N,n,(n-1)/2}(p) - L_{N+1,n,(n-1)/2}(p) \begin{cases} = 0 & \text{für} \quad 0 \leqq p \leqq \dfrac{n-1}{2(N+1)} \\ > 0 & \text{für} \quad \dfrac{n-1}{2(N+1)} < p < \dfrac{1}{2} \\ = 0 & \text{für} \quad p = \dfrac{1}{2} \\ < 0 & \text{für} \quad \dfrac{1}{2} < p < 1 - \dfrac{n-1}{2(N+1)} \\ = 0 & \text{für} \quad 1 - \dfrac{n-1}{2(N+1)} \leqq p \leqq 1. \end{cases}$$

Wie schon das Beispiel $n = 4$, $c = 1$ zeigt, hängt der Schnittpunkt p_0 in den noch verbleibenden Fällen von N ab. Um auch hier Monotonie-Aussagen machen zu können, wollen wir die Differenz $\tilde{L}_{N,n,c}(p) - \tilde{L}_{N+1,n,c}(p)$ geeignet umformen. Es ist für $0 \leqq c < n \leqq N$ und alle reellen p

$$\binom{N}{n}\{\tilde{L}_{N,n,c}(p) - \tilde{L}_{N+1,n,c}(p)\} =$$

$$= \sum_{m=0}^{c} \binom{pN}{m}\binom{N-pN}{n-m} - \sum_{m=0}^{c} \binom{pN+p}{m}\binom{N-pN-p}{n-m} +$$

$$+ \sum_{m=0}^{c} \binom{pN+p}{m}\binom{N-pN-p}{n-m} -$$

$$- \frac{N-n+1}{N+1} \sum_{m=0}^{c} \binom{pN+p}{m}\binom{N-pN-p+1}{n-m} =$$

$$= \binom{N}{n}\left\{\tilde{L}_{N,n,c}(p) - \tilde{L}_{N,n,c}\left(p + \frac{p}{N}\right)\right\} +$$

$$+ \sum_{m=0}^{c} \binom{pN+p}{m}\binom{N-pN-p}{n-m-1}\frac{m-pn}{n-m}. \tag{13}$$

Die rechts stehende Summe läßt sich noch weiter vereinfachen, und zwar beweisen wir durch vollständige Induktion nach c:

$$\sum_{m=0}^{c} \binom{pN+p}{m}\binom{N-pN-p}{n-m-1}\frac{m-pn}{n-m} =$$

$$= -p\binom{pN+p-1}{c}\binom{N-pN-p}{n-c-1}. \tag{14}$$

Ersichtlich ist (14) für $c = 0$ richtig, und der Schluß von $(c-1)$ auf c, wobei $1 \leqslant c \leqslant n-1$ ist, ergibt sich durch elementare Rechnung:

$$\sum_{m=0}^{c} \binom{p\,N+p}{m}\binom{N-p\,N-p}{n-m-1}\frac{m-p\,n}{n-m} =$$

$$= -p\binom{p\,N+p-1}{c-1}\binom{N-p\,N-p}{n-c} +$$

$$+\binom{p\,N+p}{c}\binom{N-p\,N-p}{n-c-1}\frac{c-p\,n}{n-c} = -\binom{p\,N+p-1}{c-1}\binom{N-p\,N-p}{n-c-1}\cdot$$

$$\cdot\left\{\frac{N-p\,N-p-n+c+1}{n-c}\,p - \frac{(p\,N+p)\,(c-p\,n)}{c(n-c)}\right\} =$$

$$= -p\binom{p\,N+p-1}{c-1}\binom{N-p\,N-p}{n-c-1}\frac{p\,N+p-c}{c} =$$

$$= -p\binom{p\,N+p-1}{c}\binom{N-p\,N-p}{n-c-1}.$$

Setzen wir (14) in (13) ein, so erhalten wir für $0 \leqslant c < n \leqslant N$ und alle reellen p:

$$\binom{N}{n}\{\bar{L}_{N,n,c}(p) - \bar{L}_{N+1,n,c}(p)\} =$$

$$= \binom{N}{n}\left\{\bar{L}_{N,n,c}(p) - \bar{L}_{N,n,c}\left(p+\frac{p}{N}\right)\right\} - p\binom{p\,N+p-1}{c}\binom{N-p\,N-p}{n-c-1}. \quad (15)$$

Gleichung (15) gestattet zusammen mit der Abschätzung des Hilfssatzes 1 die gewünschte Monotonie-Aussage:

Satz 4: Es sei $0 < c < n-1 < N$. Dann ist

$$L_{N,n,c}(p) - L_{N+1,n,c}(p)\begin{cases} = 0 & \text{für} \quad 0 \leqslant p \leqslant \dfrac{c}{N+1} \\ > 0 & \text{für} \quad \dfrac{c}{N+1} < p \leqslant \dfrac{c}{n-1} - \dfrac{c}{(n-1)\,(N+1)} \\ < 0 & \text{für} \quad \dfrac{c}{n-1} + \dfrac{n-c-1}{(n-1)\,(N+1)} \leqslant p < 1 - \dfrac{n-c-1}{N+1} \\ = 0 & \text{für} \quad 1 - \dfrac{n-c-1}{N+1} \leqslant p \leqslant 1. \end{cases}$$

Beweis: Aus (2) und Satz 1 folgt die Behauptung für

$$0 \leqslant p \leqslant \frac{c}{N} \qquad \text{und} \qquad 1 - \frac{n-c-1}{N} \leqslant p \leqslant 1.$$

Für die verbleibenden p können wir L jeweils durch $\tilde{L}$ ersetzen. Wir schreiben nun (15) in der Form

$$\binom{N}{n}\{\tilde{L}_{N,n,c}(p) - \tilde{L}_{N+1,n,c}(p)\} =$$

$$= \binom{N}{n}\left\{\tilde{L}_{N,n,c}\left(\frac{N+1}{N}p - \frac{p}{N}\right) - L_{N,n,c}\left(\frac{N+1}{N}p\right)\right\} -$$

$$- p\binom{pN+p-1}{c}\binom{N-pN-p}{n-c-1},$$

wobei wir die rechte Seite nach Hilfssatz 1 abschätzen können. Nach (11) ist

$$\binom{N}{n}\left\{\tilde{L}_{N,n,c}\left(\frac{N+1}{N}p - \frac{p}{N}\right) - \tilde{L}_{N,n,c}\left(\frac{N+1}{N}p\right)\right\} >$$

$$> p\binom{pN+p-1}{c}\binom{N-pN-p}{n-c-1}$$

für

$$\frac{c}{N} \leqslant \frac{N+1}{N}p \leqslant \frac{c}{n-1},$$

also für

$$\frac{c}{N+1} \leqslant p \leqslant \frac{c}{n-1} - \frac{c}{(n-1)(N+1)}.$$

Für

$$\frac{c}{n-1} + \frac{1}{N} \leqslant \frac{N+1}{N}p \leqslant 1 - \frac{n-c-2}{N},$$

also für

$$\frac{c}{n-1} + \frac{n-c-1}{(n-1)(N+1)} \leqslant p \leqslant 1 - \frac{n-c-1}{N+1}$$

folgt aus (12)

$$\binom{N}{n}\left\{L_{N,n,c}\left(\frac{N+1}{N}p-\frac{p}{N}\right)-L_{N,n,c}\left(\frac{N+1}{N}p\right)\right\}<$$
$$<p\binom{pN+p-1}{c}\binom{N-pN-p}{n-c-1}.$$

5. Vergleich der Operations-Charakteristiken

Unabhängig von den Ergebnissen des vorigen Abschnitts lassen sich die Operations-Charakteristiken $L_{N,n,c}(p)$ und $L_{n,c}(p)$ miteinander vergleichen, wenn N sehr groß ist. Man betrachtet dazu die Hilfsfunktion

$$\lambda(N,n,c,p)=N(N-1)\cdot\ldots\cdot(N-n+1)\{L_{N,n,c}(p)-L_{n,c}(p)\}.$$

Sie ist als Funktion von N ersichtlich ein Polynom vom Grad $(n-1)$. Durch vollständige Induktion nach c läßt sich leicht beweisen, daß der Koeffizient von N^{n-1} für $c=0,1,\ldots,n-1$ lautet:

$$\frac{n}{2(n-1)}\binom{n-1}{c}p^c(1-p)^{n-c-1}\left(\frac{c}{n-1}-p\right).$$

Für „große" N ist also das Vorzeichen von $L_{N,n,c}(p)-L_{n,c}(p)$ gleich dem Vorzeichen von $(c/(n-1)-p)$.

Genauere Aussagen lassen sich mit Hilfe des vorigen Abschnittes machen, wobei stets noch berücksichtigt werden muß, daß $L_{n,c}(p)$ als Funktion von p bekanntlich streng monoton abnehmend ist.

Für $c=0$ und $c=n-1$ folgt aus Satz 2 unmittelbar die in der Einleitung zitierte Eigenschaft (4).

Aus Satz 4 folgt sofort

Satz 5: Für $0<c<n-1<N$ *ist*

$$L_{N,n,c}(p)-L_{n,c}(p)\begin{cases}=0 & \text{für} \quad p=0\\ >0 & \text{für} \quad 0<p\leqq\dfrac{c}{n-1}-\dfrac{c}{(n-1)(N+1)}\\ <0 & \text{für} \quad \dfrac{c}{n-1}+\dfrac{n-c-1}{(n-1)(N+1)}\leqq p<1\\ =0 & \text{für} \quad p=1.\end{cases}$$

Wenn n ungerade und $c=(n-1)/2$ ist, läßt sich diese Aussage mit Hilfe von Satz 3 noch verschärfen:

$$L_{N,n,(n-1)/2}(p) - L_{n,(n-1)/2}(p) \begin{cases} > 0 & \text{für} \quad 0 < p < \frac{1}{2} \\ = \frac{1}{2} - \frac{1}{2} = 0 & \text{für} \quad p = \frac{1}{2} \\ < 0 & \text{für} \quad \frac{1}{2} < p < 1. \end{cases} \tag{16}$$

Zur Beantwortung der in der Einleitung im Zusammenhang mit der Forderung (3) aufgeworfenen Frage genügt die allgemein gültige, aber schwächere Aussage des Satzes 5, denn wie dort bereits erwähnt ist, ist $c/(n-1)$ ungefähr gleich dem 50%-Punkt von $L_{n,c}(p)$. Genauer noch läßt sich zeigen:

Satz 6: Für die Operations-Charakteristik $L_{n,c}(p)$ mit $n \geqslant 2$ gilt:

$$L_{n,c}\left(\frac{c}{n-1}\right) > \frac{1}{2} > L_{n,c}\left(\frac{c+1}{n+1}\right) \qquad \text{falls} \qquad 0 \leqslant c < \frac{n-1}{2}$$

$$L_{n,c}\left(\frac{c}{n-1}\right) = \frac{1}{2} = L_{n,c}\left(\frac{c+1}{n+1}\right) \qquad \text{falls} \qquad c = \frac{n-1}{2}$$

$$L_{n,c}\left(\frac{c+1}{n+1}\right) > \frac{1}{2} > L_{n,c}\left(\frac{c}{n-1}\right) \qquad \text{falls} \qquad \frac{n-1}{2} < c \leqslant n-1.$$

Beweis: Für $c = (n-1)/2$ ist $c/(n-1) = (c+1)/(n+1) = 1/2$ und die 2. Behauptung sofort einzusehen. Für $c \neq (n-1)/2$ ziehen wir den Aufsatz [2] heran, in dem der Zentralwert der k-ten Ranggröße $\xi^{(k)}$ einer Stichprobe vom Umfang $n \geqslant 2$ untersucht wird, und zwar wird gezeigt, daß dieser Zentralwert zwischen dem Quantil $z((k-1)/(n-1))$ der Ordnung $(k-1)/(n-1)$ und dem Quantil $z(k/(n+1))$ der Ordnung $k/(n+1)$ liegt. Im einzelnen besagt der dortige Satz 5 zusammen mit der dortigen Gleichung (2): Für $n \geqslant 2$, $k < (n+1)/2$ ist

$$W\left(\xi^{(k)} \leqslant z\left(\frac{k-1}{n-1}\right)\right) = 1 - \sum_{m=0}^{k-1} \binom{n}{m} \left(\frac{k-1}{n-1}\right)^m \left(1 - \frac{k-1}{n-1}\right)^{n-m} <$$

$$< \frac{1}{2} < W\left(\xi^{(k)} \leqslant z\left(\frac{k}{n+1}\right)\right) = 1 - \sum_{m=0}^{k-1} \binom{n}{m} \left(\frac{k}{n+1}\right)^m \left(1 - \frac{k}{n+1}\right)^{n-m}$$

und für $n \geqslant 2$, $k > (n+1)/2$ ist

$$W\left(\xi^{(k)} \leqslant z\left(\frac{k}{n+1}\right)\right) = 1 - \sum_{m=0}^{k-1} \binom{n}{m} \left(\frac{k}{n+1}\right)^m \left(1 - \frac{k}{n+1}\right)^{n-m} <$$

$$< \frac{1}{2} < W\left(\xi^{(k)} \leqslant z\left(\frac{k-1}{n-1}\right)\right) = 1 - \sum_{m=0}^{k-1} \binom{n}{m} \left(\frac{k-1}{n-1}\right)^m \left(1 - \frac{k-1}{n-1}\right)^{n-m}.$$

Setzen wir $c = k-1$, so ergibt sich sofort die 1. und die 3. Behauptung unseres Satzes 6.

Da in der statistischen Qualitätskontrolle die Wahrscheinlichkeit p — als Ausschußanteil — im allgemeinen klein ist, spielt dort die Approximation der Binomial-Verteilung durch die Poisson-Verteilung eine besonders wichtige Rolle. In Ergänzung zu Satz 5 wollen wir daher die Operations-Charakteristik $L_{n,c}(p)$ auch noch mit der Näherung

$$L^*_{n,c}(p) = \sum_{m=0}^{c} \frac{(np)^m}{m!} e^{-np}$$

vergleichen, die sich aus der Approximation durch die Poisson-Verteilung ergibt.

Satz 7: Für $0 \leqq c < n - 1$ *und* $0 < p \leqq 1$ *ist*

$$L_{n,c}(p) < L^*_{n,c}(p).$$

Anmerkung: Wie schon das Beispiel $c = 1$, $n = 2$ zeigt, ist der Satz für $c = n - 1$ nicht richtig.

Beweis: Es ist

$$\frac{d}{dp} L_{n,c}(p) = \frac{-n!}{c!(n-c-1)!} p^c (1-p)^{n-c-1}$$

$$\frac{d}{dp} L^*_{n,c}(p) = \frac{-n^{c+1} p^c}{c!} e^{-np}$$

$$L_{n,c}(0) - L^*_{n,c}(0) = 0, \qquad \frac{d}{dp}\{L_{n,c}(p) - L^*_{n,c}(p)\}|_{p=0} = 0$$

$$L_{n,c}(1) - L^*_{n,c}(1) = -L^*_{n,c}(1) < 0, \qquad \frac{d}{dp}\{L_{n,c}(p) - L^*_{n,c}(p)\}|_{p=1} > 0.$$

Da

$$\frac{d}{dp}\{L_{n,c}(p) - L^*_{n,c}(p)\}$$

ersichtlich höchstens eine Nullstelle im offenen Intervall (0, 1) hat, kann es keine Stelle p_1 mit $0 < p_1 < 1$ und $L_{n,c}(p_1) - L^*_{n,c}(p_1) > 0$ geben.

Literatur

[1] FELLER, W.: An Introduction to Probability Theory and Its Applications. Vol. I, Second Edition, Wiley, New York, London, Sydney, 1957.
[2] UHLMANN, W.: Ranggrößen als Schätzfunktionen. Metrika 7 (1963) 23—40.
[3] UHLMANN, W.: Statistische Qualitätskontrolle. Teubner-Verlag, Stuttgart, 1966.