DER LIKELIHOOD-QUOTIENT ZUR ER - FASSUNG DES SUBJEKTIVEN SIGNIFIKANZNIVEAUS

von

Jürgen KRIZ

Forschungsbericht No. 9

Juli 1967

Einleitung

In den letzten 15 Jahren haben die Begriffe "subjektive Wahrscheinlichkeit" und "subjektives Signifikanzniveau" immer mehr Bedeutung in der Psychologie - insbesondere auf dem Gebiet der Entscheidungstheorie - gewonnen.
Besonders eit der Konzipierung des SEU-Modells durch Ward EDWARDS (1955) sind zahlreiche Arbeiten vorwiegend von Psychologen, aber auch von Ökonomen, Sozialwissenschaftlern und Mathematikern entstanden, die sich mit der subjektiven Wahrscheinlichkeit teils theoretisch-methodisch, teils praktisch-empirisch auseinandersetzen. Schon 1964 war die Zahl dieser Arbeiten so angewachsen, dass EDWARDS in seine Bibliographie (EDWARDS 1964) über 700 Titel aufnehmen musste, und es - laut EDWARDS - leicht gefallen wäre, eine doppelt so lange Liste zu erstellen.
Die vorliegende Arbeit will daher gar nicht erst den Versuch unternehmen, auch nur die wesentlichsten Arbeiten der letzten Jahre zu referieren; sondern es müssen die wichtigsten Arbeiten und Begriffe als bekannt vorausgesetzt werden. Verwiesen sei dabei vielleicht auf die sehr guten Sammelreferate von EDWARDS (1954 und 1961), in denen das Wesentlichste bis April 1960 zusammengefasst ist, sowie auf Charles VLEK, der im ersten Teil einer Arbeit (VLEK 1965) ebenfalls einige wichtige Arbeiten kurz referiert.

Das Verfahren von WENDT

1966 wurde von Dirk WENDT ein Verfahren zur Erfassung des subjektiven Signifikanzniveaus[1] vorgeschlagen, das im wesentlichen auf die Sequenzanalyse von Abraham WALD (1947) zurückgeht.

In der Sequenzanalyse für Alternativmerkmale kann mit Hilfe des sogenannten "Likelihood-Quotienten" zwischen zwei Alternativhypothesen H_1 und H_2 entschieden werden, die hinsichtlich der Verteilung von Alternativdaten in einer Grundgesamtheit (z.B. ein bestimmtes Mischungsverhältnis von gelben und grünen Kugeln in einem Beutel) bestehen. Dabei ist der Likelihood-Quotient die Wahrscheinlichkeit der empirisch beobachteten Daten unter der Hypothese H_1 dividiert durch die Wahrscheinlichkeit dieser Daten unter der Hypothese H_2:

$$L = \frac{P(D/H_1)}{P(D/H_2)} \qquad (1)$$

Bestehen nun folgende Hypothesen über die Wahrscheinlichkeit für das Auftreten von Merkmal A (bzw. B) in der Grundgesamtheit:

$H_1: P(A) = \pi_1$ bzw. $P(B) = 1 - \pi_1$ und

$H_2: P(A) = \pi_2$ bzw. $P(B) = 1 - \pi_2$, so gilt bei einer Abfolge von x Alternativen der einen und y Alternativen der anderen Art: (insgesamt $N = x+y$ Beobachtungen)

$$L = \frac{\binom{x+y}{x} \pi_1^y (1 - \pi_1)^x}{\binom{x+y}{x} \pi_2^y (1 - \pi_2)^x} \qquad (y = N - x)$$

$$= \left(\frac{\pi_1}{\pi_2}\right)^y \left(\frac{1 - \pi_1}{1 - \pi_2}\right)^x \qquad (2)$$

---

[1] WENDT spricht von "subjektivem Verlässlichkeitsniveau" (SVN); - dennoch soll die aus einer früheren Arbeit stammende Bezeichnung "subjektives Signifikanzniveau" in dieser Arbeit beibehalten werden.

Sofern nun L einen bestimmten Wert $A = (1 - \beta) / \alpha$ erreicht oder überschreitet, wird H_1 mit den Irrtumswahrscheinlichkeiten α und β angenommen; sofern L einen anderen Wert $B = \beta / (1 - \alpha)$ erreicht oder unterschreitet, wird H_2 mit denselben Irrtumswahrscheinlichkeiten angenommen. Solange L keine dieser beiden Bedingungen erfüllt (also $A \gg L \gg B$), kann noch keine Entscheidung getroffen, sondern es muss weiterhin Information (d.h. weitere Daten) erhoben werden.[1)]

In dem von WENDT vorgeschlagenen Verfahren wird dieser Entscheidungsprozess umgekehrt. Der Grundgedanke dabei ist, dass (genau wie aus dem vorgegebenen α und β das A und B für die Entscheidung berechnet werden), aus dem A und B nach einer erfolgten Entscheidung das α und β, das diesem Entscheidungsprozess zugrunde gelegen hat, rückrechenbar sein müsste. Dazu wird eine Vp in eine konkrete Entscheidungssituation gestellt, in der sie sich zwischen zwei Hypothesen über die Verteilung von Alternativmerkmalen in einer Grundgesamtheit zu entscheiden hat, wobei versichert wird, dass eine der beiden Hypothesen den wahren Sachverhalt (π_o)[2)] garantiert richtig darstellt.

In den WENDTschen Versuchen wurde z.B. Vpn ein Beutel mit der Versicherung gezeigt, dass dieser entweder 70% rote und 30% gelbe oder 30% rote und 70% gelbe Kugeln enthalte. Die Vpn konnten nun beliebig oft je eine Kugel aus dem Beutel entnehmen (und zurücklegen, damit π_o konstant blieb), wobei jede Entnahme mit bestimmten "Kosten" (Punkte, die später in Geld umgerechnet wurden) verbunden war. Ferner wurden richtige und falsche Entscheidungen für H_1 bzw. H_2 "belohnt" oder "bestraft", (ebenfalls in Form von Punkten), so dass die Entscheidungssituation durch eine Payoff-Matrix charakterisiert wird.[3)]

---

1) genaueres siehe WALD 1947

2) also entweder π_1 oder π_2

3) näheres siehe WENDT 1966

Wenn man die Hypothese, für die sich eine Vp entschieden hat, mit H_1 bezeichnet, kann man nun ihr Verhalten so interpretieren, dass vor der Entscheidung ihr "subjektiver Likelihood-Quotient" unter der kritischen Grenze gewesen sein muss, (d.h. $L'_{x+y-1} < A$) und erst durch die Information der letzten Kugel diese Grenze erreicht oder überschritten wurde (d.h. $L'_{x+y} \geqq A$).

Da sich die Vp nur für eine Hypothese entscheidet, (die mit H_1 bezeichnet wurde) kann nur eine Art der Irrtumswahrscheinlichkeit bestimmt werden. Da aus $A = (1 - \beta')/\alpha'$ das α' nicht ohne Kenntnis von β' bestimmt werden könnte, werden α' und β' gleichgesetzt.

Daraus folgt wiederum, dass das Interesse der Vpn, sich für eine Hypothese zu entscheiden, bei beiden Hypothesen gleich gross sein sollte, was bedeutet, dass die Payoff-Matrix gleiche gewogene Zeilenmittel haben muss. [1)]

Unter diesen Voraussetzungen gilt für α' :

$$\alpha' = \frac{1}{(A+1)} \qquad (3)$$

und WENDT kommt zu dem Ungleichungssystem:

$$1/(A(N, X_N, \pi_1, \pi_2) + 1) \leq SVN <$$

$$< 1/(A(N-1, X_{N-1}, \pi_1, \pi_2) + 1) \qquad (4)$$

wobei:

$A(N, X_N, \pi_1, \pi_2)$ der Likelihood-Quotient nach Entnahme der letzten Kugel[2)]

$A(N-1, X_{N-1}, \pi_1, \pi_2)$ der Likelihood-Quotient nach Entnahme der vorletzten Kugel

SVN das subjektive Signifikanzniveau bedeuten.

[1)] genaueres siehe WENDT 1966

[2)] Obwohl WENDT bei (3) und (4) "A' " schreibt, sei dieser Wert hier mit "A" bezeichnet, da es sich offensichtlich um <u>objektive</u> Werte handelt, nicht um subjektive.

Da sich nun die beiden Werte für A leicht nach Formel (2) berechnen lassen, erhält man somit eine obere und eine untere Schranke, zwischen denen das α' liegen muss. Das Verfahren gestattet somit zwar keine exakte Bestimmung des subjektiven Signifikanzniveaus[1] ($= \alpha'$), aber die beiden Schranken sind in der Praxis genügend nahe beieinander, so dass das α' hinreichend determiniert ist.

Betrachten wir noch einmal die 3 wichtigsten Voraussetzungen, die für dieses Verfahren gemacht wurden:

1.) $\alpha' = \beta'$

2.) $L'_{x+y-1} < A \leq L'_{x+y}$

3.) $\alpha' = f(L)$

Voraussetzung 1) besagt nur, dass die Entscheidungssituation so geartet sein muss, dass keine der Hypothesen aus irgendeinem Grunde von der Vp von vornherein bevorzugt werden darf - eine Forderung, wie sie sich - zumindest ähnlich - in fast jedem psychologischen Experiment stellt, und die durch die Beschränkung, nur Payoff-Matrizen mit gleichen gewogenen Zeilenmitteln zu verwenden, sicherlich hinreichend erfüllt wird.

Voraussetzung 2) fordert nichts anderes, als dass sich die Vp in der Entscheidungssituation zumindest subjektiv "sinnvoll" verhält, d.h. dass tatsächlich die Information der letzten Kugel die Entscheidung herbeigeführt hat und die Vp sich <u>nach</u> der Entscheidung der Richtigkeit ihrer Hypothese sicherer ist, als <u>vor</u> der Entscheidung, als sie noch weniger Information hatte. Eine Forderung, die sicher ebenfalls erfüllt wird, da man sonst annehmen müsste, dass eine Entscheidung ausschliesslich nach dem Zufall erfolgt.

---

[1] Wegen der Länge der Worte, wird im folgenden Text manchmal "subjektives Signifikanzniveau" mit SSN
"Likelihood-Quotient" mit LQ abgekürzt.

In der Voraussetzung 3) wird gefordert, dass das subjektive Signifikanzniveau eine Funktion des Likelihood-Quotienten sein muss. Obwohl WENDT diese Forderung nirgends explizit stellt, geht sie doch als Voraussetzung in das Verfahren ein, denn es ist offenbar nur sinnvoll und zulässig eine Variable (wie das subjektive Signifikanzniveau) durch eine andere Variable (wie der auf Grund experimenteller Daten gewonnene objektive Likelihood-Quotient) zu beschreiben, und sogar zu berechnen, wenn man eine funktionale Beziehung zwischen beiden postuliert.

Diese Forderung scheint ebenfalls als wenig problematisch beurteilt zu werden. Jedenfalls wird in fast allen Untersuchungen zur subjektiven Wahrscheinlichkeit eine Beziehung zwischen subjektiver und objektiver Wahrscheinlichkeit vorausgesetzt. Auch WENDT erörtert nicht, ob es überhaupt zulässig ist, das subjektive Signifikanzniveau durch den Likelihood-Quotienten zu beschreiben.

In einer früheren Arbeit gelang es dem Autor jedoch, die Abhängigkeit der subjektiven Wahrscheinlichkeit und des subjektiven Signifikanzniveaus von der objektiven Wahrscheinlichkeit zumindest für einen sehr kleinen Entscheidungsbereich in Zweifel zu stellen (KRIZ 1967). In dieser Arbeit konnte nachgewiesen werden, dass sich Vpn in einer Entscheidungssituation nicht unbedingt rational verhalten. Bei den Versuchen erwies sich ein für die objektive Wahrscheinlichkeit sehr effizienter Parameter - nämlich die Stichprobengrösse - als subjektiv völlig irrelevant.

Man kann sich nun zumindest fragen, ob es nicht möglich wäre, dass sich die Vpn bei den Versuchen von WENDT gar nicht analog zur Sequenzanalyse verhalten haben, sondern eventuell die Entscheidung durch andere Parameter beeinflusst wurden als durch den Likelihood-Quotienten.

Beitrachten wir eine spezielle Versuchssituation: Angenommen der Vp werden 2 Hypothesen über das Mischungsverhältnis von gelben und roten Kugeln in einem Beutel zur Wahl gestellt (von der garantiert eine richtig ist), so ergibt sich für den speziellen Fall, dass $\pi_2 = 1 - \pi_1$ ist (z.B. H_1: 30 rote, 70 gelbe H_2: 70 rote, 30 gelbe) aus Gleichung (2), nachdem y gelbe und x rote Kugeln aufgetreten sind ($y > x$):

$$L = \left(\frac{\pi_1}{\pi_2}\right)^y \left(\frac{1-\pi_1}{1-\pi_2}\right)^x = \left(\frac{\pi_1}{1-\pi_1}\right)^y \left(\frac{1-\pi_1}{\pi_1}\right)^x =$$

$$= \left(\frac{\pi_1}{1-\pi_1}\right)^{y-x} \qquad (5)$$

Damit ist L vollständig durch die Differenz von gelben und roten Kugeln ($y-x$) determiniert, d.h. falls nach einem bestimmten x_o und y_o die Entscheidung für H_1 getroffen wurde, so ist die Irrtumswahrscheinlichkeit die gleiche, wie wenn man sich nach x_o+n und y_o+n entschieden hätte. Ein konkretes Beispiel: Eine Vp, die sich nach 6 gelben und 3 roten Kugeln für H_1 entscheidet, müsste sich mit der selben Sicherheit nach 13 gelben und 10 roten, 23 gelben und 20 roten oder - um es noch extremer darzustellen - nach 1003 gelben und 1000 roten Kugeln für H_1 entscheiden. Obwohl dies normativ zweifellos die richtige Verhaltensweise wäre, ist es doch eine etwas unerwartete Konsequenz aus dem Modell, worauf u. a. bereits VLEK (1965) hinweist.

Stellen wir uns in einem Gedankenexperiment vor, es wären einmal 6 gelbe und 3 rote Kugeln gezogen worden (Fall 1), in einem anderen Fall (Fall 2) wären 13 gelbe und 10 rote Kugeln aufgetreten (die letzte Kugel war jeweils eine gelbe). Würden wir uns wirklich in beiden Fällen mit gleicher Sicherheit für H_1 entschieden haben?

Sofern wir dies nicht annehmen, denken wir uns einen Fall 3, bei dem 10 rote Kugeln aufgetreten sind und so viel gelbe, dass (nachdem die letzte eine gelbe war) wir uns mit derselben Irrtumswahrscheinlichkeit wie in Fall 1 entscheiden konnten.

Es würde dann subjektiv gelten:

1.) Bei gleichen y-x in Fall 1 und 2 wurde mit verschiedenen Irrtumswahrscheinlichkeiten entschieden.
2.) Gleiche Irrtumswahrscheinlichkeit liegt in Fall 1 und 3 nur dann vor, wenn $(y-x)_{Fall_1} \neq (y-x)_{Fall_3}$ ist.

Wenn man diese drei Fälle aber mit dem WENDTschen Verfahren untersuchte, käme man zu folgenden Schlüssen:

1.) $(y-x)_{Fall_1} = (y-x)_{Fall_2}$, d.h. $\alpha_1 = \alpha_2 \Rightarrow \alpha'_1 = \alpha'_2$,

d.h. es wurde in Fall 1 und 2 mit derselben Irrtumswahrscheinlichkeit entschieden.

2.) $(y-x)_{Fall_1} \neq (y-x)_{Fall_3}$, d.h. $\alpha_1 \neq \alpha_2 \Rightarrow \alpha'_1 \neq \alpha'_2$,

d.h. es wurde in Fall 1 und 3 mit verschiedenen Irrtumswahrscheinlichkeiten entschieden.

Es ergibt sich also, dass bei den 3 Gedankenexperimenten das Verhalten der Vp falsch interpretiert werden würde. Dies zeigt zumindest, dass die Annahme, das SSN liesse sich durch den LQ beschreiben, nicht so unproblematisch ist, wie es Anfangs zu sein schien, und eine nähere Untersuchung dieser Annahme scheint gerechtfertigt.

Die Fragestellung dieser Arbeit lautet also:

Fällen Vpn wirklich im Sinne der Sequenzanalyse ihre Entscheidungen, d.h. ist der Likelihood-Quotient wirklich ein adäquates Mass, das subjektive Signifikanzniveau zu erfassen?

Für die weitere Bezugnahme wird formuliert:

H_o: Die Entscheidung erfolgt so, dass das SSN mit dem Likelihood-Quotienten erfasst und beschrieben werden kann.

H_A: Das SSN lässt sich mit Hilfe des LQ nicht erfassen., Die Berechnung des SSN über den objektiven LQ würde zu Werten führen, die nicht einmal in eindeutiger Ordinalrelation zueinander stehen (vergl. Gedankenexperiment).

Versuchsplan

Der Versuchsplan baut auf dem beschriebenen Gedankenexperiment auf. Setzt man nach Vorr. 2 $\alpha = \beta$, gilt :

$$H_1 \text{ wenn } \quad \frac{1-\alpha}{\alpha} \leqq \left(\frac{\pi_1}{\pi_2}\right)^y \left(\frac{1-\pi_1}{1-\pi_2}\right)^x \tag{6a}$$

$$H_2 \text{ wenn } \quad \frac{\alpha}{1-\alpha} \geqq \left(\frac{\pi_1}{\pi_2}\right)^y \left(\frac{1-\pi_1}{1-\pi_2}\right)^x \tag{6b}$$

Die Ungleichungen durch Gleichungen ersetzt, logarithmiert und nach y aufgelöst:

$$H_1 \text{ wenn } \quad y = -x \left(\frac{\log \frac{1-\pi_1}{1-\pi_2}}{\log \frac{\pi_1}{\pi_2}}\right) + \frac{\log \frac{1-\alpha}{\alpha}}{\log \frac{\pi_1}{\pi_2}} \tag{7a}$$

$$H_2 \text{ wenn } \quad y = -x \left(\frac{\log \frac{1-\pi_1}{1-\pi_2}}{\log \frac{\pi_1}{\pi_2}}\right) - \frac{\log \frac{1-\alpha}{\alpha}}{\log \frac{\pi_1}{\pi_2}} \tag{7b}$$

$$\left(\text{denn } \log \frac{\alpha}{1-\alpha} = -\log \frac{1-\alpha}{\alpha}\right)$$

Es zeigt sich die bekannte Tatsache, dass 7a und 7b zwei Geradengleichungen der Form $y = ax + b$ darstellen, die sich nur in "b" unterscheiden und somit parallel laufen. Für den Fall $\alpha = \beta$ ist es sinnvoll, noch eine dritte Gerade, die die Form

$$y = -x \left(\frac{\log \frac{1-\pi_1}{1-\pi_2}}{\log \frac{\pi_1}{\pi_2}}\right)$$

hat, zu betrachten. Diese Gerade ist offenbar von den beiden "Entscheidungsgeraden" gleichweit entfernt und repräsentiert die Punkte höchster Indifferenz zwischen H_1 und H_2.

Sofern die WENDTschen Annahmen richtig sind, sollte eine Vp sich von einem Punkt höchster Indifferenz immer nach einer konstanten Häufigkeit (($\log \frac{1-\alpha}{\alpha}$)/($\log \frac{\pi_1}{\pi_2}$)) des Auftretens einer Alternative für H_1 (bzw. H_2) entscheiden, unabhängig davon, wo dieser Punkt auf der Geraden liegt, d.h. unabhängig davon, wie gross die Stichprobe vorher war.

Obwohl dies allgemein für jedes Hypothesenpaar H_1 und H_2 gilt, wurde für die Experimente aus Gründen der Einfachheit der Sonderfall (vergl. S. 7) gewählt, bei dem die beiden Hypothesen symmetrisch sind (d.h. es gilt $\pi_1 + \pi_2 = 1$)

Aus 7a und 7b wird dann nämlich:

$$y = x \pm (\log \frac{1-\alpha}{\alpha})/(\log \frac{\pi_1}{\pi_2})^{1)} \qquad (8)$$

Alle drei Geraden haben hier den Anstieg "+1" (= 45°) und die "Indifferenzgerade" stellt die Verteilung n_{gelbe} / n_{rote} Kugeln bzw. P_{indiff} = 0.5 dar. Diese Vereinfachung erschien zulässig, da ein eventueller Fehler nur die Verifizierung der Arbeitshypothese erschweren würde.

Die Versuche wurden analog zu den Versuchen von WENDT durchgeführt. Den Vpn wurden 2 völlig gleiche Beutel (von denen die Vp einen wählte) mit folgender Instruktion gezeigt :

"In diesem Beutel sind genau 100 Spielsteine. Und zwar entweder 60 gelbe und 40 grüne, oder 40 gelbe und 60 grüne.(Eine andere Mischung ist nicht möglich). Ich werde aus diesem Beutel nacheinander zufällig Spielsteine ziehen. Ihre Aufgabe ist es, auf Grund dieser Stichprobe das wahre Verhältnis in dem Beutel zu ermitteln ..... Sie können so viele Steine ziehen lassen, wie Sie wollen ....."

Im weiteren Verlauf der Instruktion wurde die Payoff-Matrix festgelegt. Von einem Ausgangskapital N wurde pro gezogenem Stein ein Punkt als Kosten abgezogen. Es war erlaubt, jederzeit aufzuhören und sich für H_1 oder H_2 zu entscheiden. War die Entscheidung richtig, erhielt die Vp zusätzlich 35 Punkte, war sie

---

1) $-x\left(\frac{\log \frac{1-\pi_1}{1-\pi_2}}{\log \frac{\pi_1}{\pi_2}}\right)$ wird für $\pi_1 + \pi_2 = 1$ zu $-x(-1) = x$

falsch, wurden 35 Punkte abgezogen (eine negative Summe war zulässig). Ein Punkt entsprach einem Schilling. (Der Betrag wurde aber nicht wirklich ausgezahlt)

Im Gegensatz zu den Versuchen von WENDT wurde aber nicht zufällig gezogen, sondern die Steine waren so präpariert, dass der Vl beim Mischen gelbe von grünen Steinen unterscheiden und somit die Folge der Steine lenken konnte.

Auf diese Weise war es möglich, nur zwei verschiedene Alternativfolgen zu ziehen :

Folge A: X Y X Y X Y Y Y Y . . . bis zur Entscheidung
Folge B: X Y X Y X Y X X Y Y X Y X Y X Y Y Y . . . b. z. Entsch.

Der Anfangsbetrag N war bei Folge A 20 Punkte, bei Folge B 30 Punkte, so dass in dem Augenblick, wo nur noch eine Farbe gezogen wurde, der Restbetrag in beiden Folgen gleich (nämlich 14 Punkte) war.

Die Vpn waren einige Kollegen und Personal am "Institut für Höhere Studien", sowie einige Psychologiestudenten - also z. T. Vpn, die Erfahrung mit statistischen Fragestellungen haben. Da aber anzunehmen ist, dass sich statistisch geschulte Vpn bei einer statistischen Entscheidungssituation eher "rational" verhalten als andere, würde ein Fehler wieder nur die Verifizierung der Arbeitshypothese erschweren. - Bei je 4 Vpn waren Y_A Y_B : GeGe, GrGr, GeGr, GrGe. Bei 8 Vpn wurde zuerst Folge A gezogen, bei 8 zuerst Folge B. Die Zeit zwischen beiden Versuchen lag zwischen wenigen Minuten und 10 Tagen.

Ergebnisse

Wenn der LQ ein adäquates Mass zur Beschreibung des SSN sein soll, muss erwartet werden, dass (y-x) bei Folge A und B für jede Vp gleich ist. Es gibt keinen Grund für die Annahme, dass sich eine Vp in Situation A mit einer anderen Sicherheit entscheidet als in Situation B. (Bei den meisten Vpn wurde beim 2. Versuch -gleich ob A oder B- darauf hingewiesen, dass sie sich möglichst mit der gleichen Sicherheit wie im 1. Versuch entscheiden sollten. Vergl. auch S. 13).

Die Ergebnisse der 16 Vpn sehen wie folgt aus :

Vp	y (bei x=3)	y, (bei x=8)	$D_1=y-x$ (bei x=3)	$D_2=y-x$ (bei x=8)	$D_2 - D_1$
1	6	13	3	5	+2
2	6	14	3	6	+3
3	15	22	12	14	+2
4	5	13	2	5	+3
5	7	13	4	5	+1
6	7	19	4	11	+7
7	6	15	3	7	+4
8	6	12	3	4	+1
9	6	13	3	5	+2
10	6	14	3	6	+3
11	9	13	6	5	-1
12	6	13	3	5	+2
13	6	11	3	3	0
14	10	20	7	12	+5
15	6	13	3	5	+2
16	5	14	2	6	+4

Wenn man von Vp 11 und 13 absieht, ist $D_2 - D_1$ positiv (der Modalwert ist 2, der Mittelwert liegt bei 3) . Es gibt nur zwei Möglichkeiten, dieses Ergebnis zu interpretieren:

1.) Sofern man H_o aufrecht erhält, müsste irgendeine Erklärung dafür gefunden werden, warum die Vpn bei einer grösseren Stichprobe auf einem höheren Verlässlichkeitsniveau entscheiden, zumal die Versuchsbedingungen eher das Gegenteil erwarten liessen (höherer Verlust und grössere Mühe der Stichprobe bei Versuchsbedingung B) . Aber selbst wenn man eine solche Erklärung finden würde, wäre die Messung des SSN mittels des LQ in Frage gestellt, denn zwei berechnete liessen sich nicht vergleichen, wenn sie aus verschieden grossen Stichproben gewonnen wurden.(Was normalerweise der Fall ist, wenn man wirklich zufällig zieht) . Man könnte nämlich dann keine Aussage darüber machen, wie weit der Unterschied zwischen $_1$ und $_2$ einen "wahren" Unterschied zwischen Vp_1 und Vp_2 (Situation 1 und 2) darstellt und wie weit er einfach auf die verschiedene Stichprobengrösse zurückzuführen ist, 2.) Die zweite - und dem Autor sinnvoller erscheinende - Möglichkeit, die Ergebnisse zu interpretieren besteht darin, das SSN als unabhängig vom objektiven LQ anzunehmen. Damit soll nicht gesagt sein, dass die objektiven Wahrscheinlichkeiten überhaupt keinen Einfluss auf die Entscheidungssituation hätten, aber die Beziehung zwischen LQ und SSN ist derart, dass sich der LQ nicht zur Beschreibung des SSN eignet, d.h. die Vpn entscheiden zwar nicht völlig irrational, aber auch nicht im Sinne der Sequenzanalyse. Dafür spricht neben den dargestellten Ergebnissen noch eine weitere Tatsache: Den Vpn wurde nach dem 2. Versuch die Versuchsanordnung bekannt gegeben, und sie wurden aufgefordert, möglichst genau die Kriterien zu beschreiben, nach denen sie sich entschieden hatten. Dabei wurde von allen 14 Vpn geäussert, dass vor allem das Verhältnis von gelben zu grünen Steinen ausschlaggebend gewesen sei, und dass es daher "selbstverständlich" sei, dass bei einer grösseren Stichprobe zwar <u>relativ</u> genau soviel gelbe, <u>absolut</u> aber "natürlich" mehr gelbe Steine zusätzlich gezogen werden müssten, bevor man genau so sicher sein könnte. Manchen Vpn war auch nach längerer Diskussion nicht klar zu machen, dass zumindest objektiv ausschliesslich die Differenz zwischen gelben und grünen Steinen, nicht aber das Verhältnis ausschlaggebend ist.[1]

Auf Grund dieser Aussagen und der Versuchsergebnisse scheint es dem Autor notwendig, das von WENDT vorgeschlagene Verhaltensmodell zu erweitern. In einer Entscheidungssituation würde dann gelten:

$$SSN = f(\,LQ, N, A_1, A_2 \ldots\ldots A_n, P\,)$$

---

[1] Nach diesen Aussagen scheinen die Vpn der (irrigen) Annahme zu unterliegen, dass statt $y-x = k$, $\frac{y}{x} = k$ (für ein festes α) gilt. Dies wäre sehr interessant, da nämlich aus

$$\frac{y}{x} = K \text{ und } y = x + \Delta x, \quad \frac{x+\Delta x}{x} = K \Rightarrow \frac{\Delta x}{x} = K', \quad (K = 1+K')$$

folgen würde, was bedeutet, dass hier für die subjektive Wahrscheinlichkeit in Abhängigkeit von der Stichprobengrösse das WEBERsche Gesetz gelten würde.

wobei:

LQ = objektiver Likelihoodquotient

N = Stichprobengrösse

$A_1 .. A_n$ = sekundäre Variablen (Variablen des eigentlichen Verhaltensmodells) wie "Stichprobenkosten","Werte in der Payoff-Matrix" usw.

P = eine Persönlichkeitsvariable bedeuten soll. [1)]

Hieraus ergeben sich notwendige Forderungen für die weitere Forschung. Sofern man den Variablenkomplex $A_1 .. A_n$ untersuchen will, sollte, (bei eindimensionalen Versuchsplänen) P konstant gehalten werden, d.h. man sollte die Aussagen zunächst intraindividuell gestalten. Die Schwierigkeiten ergeben sich nun insofern, als man irgendein adäquates Mass für die SSNs braucht, um überhaupt zu Aussagen zu kommen. Nachdem gezeigt wurde, dass der LQ allein kein solches Mass ist, käme vielleicht eine Verknüpfung aus LQ und N in Frage. Über die Art dieser Verknüpfung vermag diese Arbeit aber nichts auszusagen. Eine sinnvolle Untersuchung der "A"- und "P"-Variablen ist derzeit also nur möglich, wenn man entweder Aussagen über die Verknüpfung von LQ und N machen kann, oder aber den LQ unter Konstanthaltung von N als vorläufige Beschreibung benutzt. Da aber eine Konstanthaltung von N kaum zu erreichen ist, sofern man wirklich empirisch, d.h. mit echten Zufallsfolgen arbeitet, sollte es die dringlichste Aufgabe der Forschung auf diesem Gebiet sein, die Beziehungen zwischen LQ und N genauer zu untersuchen.

Zusammenfassung und Diskussion

Es wurden die Voraussetzungen für das von WENDT 1966 vorgeschlagene Verfahren, das SSN über den objektiven LQ zu berechnen,

---

1) Da sich auch in diesem Experimet grosse interindividuelle Schwankungen ergeben haben, kann WENDTs Vorschlag, das SSN als Persönlichkeitsvariable aufzufassen, nur zugestimmt werden - zumal die Unterschiede nicht durch verschiedene Stichprobengrössen hervorgerufen sein können.

näher untersucht. Auf Grund eines Gedankenexperimentes und den Ergebnissen einer früheren Arbeit des Autors wurde die Arbeitshypothese aufgestellt, das sich der LQ möglicherweise deshalb nicht zur Beschreibung der SSN eignet, weil sich die Vpn nicht im Sinne der Sequenzanalyse verhalten (d.h. objektiv relevante Parameter subjektiv irrelevant sind und umgekehrt).

Von der Tatsache ausgehend, dass bei "symmetrischen" Hypothesen der LQ ausschliesslich von der Differenz y-x abhängt [1)] - worauf auch andere Autoren mehrfach hinwiesen, und die subjektive Gültigkeit z. T. ebenfalls in Frage stellten (PRUITT 1961, EDWARDS 1964, VLEK 1965) - wurde der Versuchsplan erstellt. Die Versuchsergebnisse zeigten, dass y-x nicht konstant blieb, sondern mit der Stichprobengrösse zunahm. Wenn man das SSN über den LQ rückgerechnet hätte, wäre die Abbildung der SSN-Skala auf die LQ-Skala nicht eindeutig. Daraus folgt, dass der LQ kein adäquates Mass sein kann, das SSN zu berechnen oder zu beschreiben.

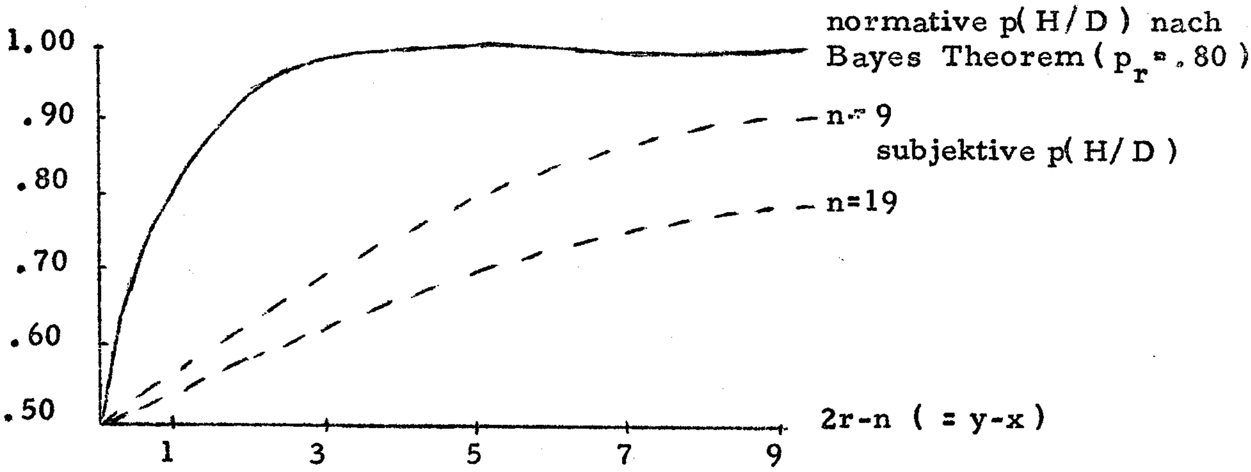
Diese Aussage ist vielleicht noch in einem anderen Bereich der Wahrscheinlichkeitsforschung interessant : In den letzten 4 Jahren haben sich zahlreiche Autoren mit der Erfassung der "posterior probabilities" (p(H/D)) beschäftigt. Zwei wesentliche Tatsachen, die von allen Autoren einheitlich festgestellt wurden, sind

1.) Vpn verhalten sich konservativ, d.h. die "posterior probabilities" werden konsistent unterschätzt.
2.) Mit zunehmender Stichprobengrösse werden die Schätzungen noch ungenauer (d.h. je grösser die Stichprobe, um so kleiner die geschätzten p(H/D)).

Die typischen Beziehungen zwischen normativen und subjektiven p (H/D) sehen so aus :

---

1) und natürlich vom Abstand der konkurrierenden Hypothesen.



"normative und subjektive p (H/D)" (aus VLEK 1965)

Sofern wir nun annehmen, dass sich die Ergebnisse dieser Arbeit auch auf den Bereich der "posterior prohabilities" übertragen lassen, d.h. dass das subjektive p (H/D) eher (oder zumindest mit) von dem Verhältnis y/x als vom LQ abhängt, würde für jedes "2r-n" das entsprechende "y/x" und somit p (H/D) mit zunehmendem Stichprobenumfang kleiner, denn

$$y / x > (y+a)/(x+a) \qquad a > 0, \quad y > x$$

was die Verhaltensweise der Vpn zumindest hinsichtlich Punkt 2 erklären würde.

EDWARDS, Ward (1954) "The theory of decision making."
Psychol. Bull. 51, 380-417

EDWARDS, Ward (1955) "The prediction of decisions among bets."
J. Exp. Psychol. 50, 201-214

EDWARDS, Ward (1961) "Behavioral decision theory."
Ann. Rev. Psychol. 12, 473-498

EDWARDS, Ward (1964) "Bibliography on decision making."
Unveröffentlichtes Manuskript

EDWARDS, Ward and PHILLIPS, Lawrence D. (1964) "Man as transducer for probabilities in Bayesion command and control systems." In: Shelly and Bryan (eds.) "Human judgments and optimality." New York: Wiley.

KRIZ, Jürgen, "Subjektives Signifikanzniveau im zwei-Stichproben-Vergleich für Alternativmerkmale."
Ztschr. f. Psychol. (vorauss. Ende 1967)

PETERSON, C.R., SCHNEIDER, R.J. and MILLER, A.J. (1965) "Sample size and the revision of subjective probabilities."
J. Exp. Psychol. 69, 522-527

PRUITT, D.G. (1961) "Informational requirements in making decisions."
Amer. J. Psychol. 74, 433-439

VLEK, Charles A.J. (1965) "The use of probabilistic information in decision making." Rep. 009-65, Psychol. Inst. Univ. of Leiden.

VLEK, Charles A.J. and BEINTEMA, Kees A. (1967) "Subjective likelihoods in posterior probability estimation."
Rep. E 014-67, Psychol. Inst. Univ. of Leiden.

WALD, Abraham (1947) "Sequential analysis." New York: Wiley.

WENDT, Dirk (1966) "Versuche zur Erfassung eines subjektiven Verlässlichkeitsniveaus." Ztschr. f. Psychol. 172, 40-81