

DER STATISTISCHE SYLLOGISMUS

1. AUSGANG

Der *Syllogismus* ist eine logische *Schluss-Form*, die erstmals von Aristoteles aufgestellt wurde. Aus der heutigen Sicht handelt es sich um *Quantoren-Logik*.

Der wohl bekannteste Syllogismus (Modus: *barbara*) hat folgende Form:

Alle G sind H	100% der G sind H	z. B. Alle Philosophen sind weise
<u>Alle F sind G</u>	<u>100% der F sind G</u>	<u>Alle Logiker sind Philosophen</u>
Alle F sind H	100% der F sind H	Alle Logiker sind weise

Es handelt sich um einen logischen Schluss der Form:

$$\text{Alle G sind H} \wedge \text{Alle F sind G} \Rightarrow \text{Alle F sind H}$$

Übersichtlicher ist eigentlich, ihn folgendermaßen zu formulieren (aber im Syllogismus wird die erste Variante gewählt):

$$\text{Alle F sind G} \wedge \text{Alle G sind H} \Rightarrow \text{Alle F sind H}$$

Alle steht bekanntlich für 100 % ($p = 1$), die beiden Prämissen und der Schluss-Satz sind also *deterministisch*, und es ist ein strenger Schluss: $p^T = 1$

Man kann als zweite These aber auch einen *individuellen* Satz wählen – obwohl Aristoteles nie individuelle Termini verwandte –, dann ergibt sich (mit Beispiel):

Alle F sind G	Alle Philosophen sind weise
<u>x_i ist ein F</u>	<u>Sokrates ist ein Philosoph</u>
x_i ist ein G	Sokrates ist weise

Beim sogenannten *Statistischen Syllogismus* steht in der ersten Prämisse statt dem *deterministischen* „alle“ ($p = 1$) ein *statistischer* Wert, zwischen 0 und 1 bzw. 0% und 100%.

Also z. B.: 80% der F sind G

Bei der zweiten Prämisse wird wiederum ein *individueller* Satz eingesetzt.

Beispiel für einen statistischen Syllogismus ist:

80% der F sind G	80% der Philosophen sind weise
<u>x_i ist ein F</u>	<u>Sokrates ist ein Philosoph</u>
x_i ist mit $p = 0,8$ ein G	Mit 80% Wahrscheinlichkeit ist Sokrates weise

Dies wäre nach dem herkömmlichen Verständnis des statistischen Syllogismus folgendermaßen zu deuten:

80% der F sind G	
<u>x_i ist ein F</u>	
<hr/>	
x_i ist ein G	0,8 Wahrscheinlichkeit

D. h. die Wahrscheinlichkeit (z. B. 0,8) wird der *logischen Ableitung* zugeschrieben, nicht dem *Schluss-Satz* ‚ x_i ist ein G‘.

2. DAS PROBLEM

Der *statistische Syllogismus* hat mich über lange Zeit beschäftigt. Ich hatte ein Modell der *graduellen logischen Folge* entwickelt, mit Hilfe der *theoretischen Wahrscheinlichkeit*. Diese Wahrscheinlichkeit nannte ich auch ‚*logische Wahrscheinlichkeit*‘, genauso wie die Schlusswahrscheinlichkeit beim statistischen Syllogismus ‚*logische Wahrscheinlichkeit*‘ genannt wird. Bei meinem Modell der logischen Wahrscheinlichkeit gab bzw. gibt es aber große Unterschiede zum statistischen Syllogismus, die ich erst nach umfangreichen Untersuchungen einordnen konnte. Um die wichtigsten Unterschiede gegenüberzustellen:

1) Modell statistischer Syllogismus

- normalerweise 2 Prämissen
- die erste Prämisse ist statistisch, sie hat eine (empirische) Wahrscheinlichkeit $0 < p < 1$
- die zweite Prämisse ist individuell (ihr wird keine Wahrscheinlichkeit zugewiesen).
- der Schluss-Satz ist individuell (ihm wird ebenfalls keine Wahrscheinlichkeit zugewiesen).
- die logische Wahrscheinlichkeit der Deduktion ist gleich dem p-Wert der ersten Prämisse

2) Modell theoretische Wahrscheinlichkeit

Dieses Modell ist viel breiter angelegt, es können eine beliebige Anzahl von Prämissen, z. B. 2 Prämissen verwendet werden, die Prämissen können deterministisch oder statistisch sein.

Die entscheidenden Unterschiede sind aber:

- allen Sätzen wird eine Wahrscheinlichkeit zugewiesen, auch den individuellen
- es muss klar unterschieden werden zwischen der *empirischen Wahrscheinlichkeit* des *Schluss-Satzes* und der *theoretischen Wahrscheinlichkeit* der Schlussfolgerung (Ableitung)
- die logische Wahrscheinlichkeit der Ableitung (p^T) ist normalerweise *ungleich* dem p-Wert der ersten Prämisse
- die Angabe der logischen Wahrscheinlichkeit verlangt i. allg. komplizierte Berechnungen und ist nicht einfach an der ersten Prämisse ablesen (wie beim statistischen Syllogismus).

Die Frage ist: Bestehen hier *zwei unterschiedliche* Arten von *logischer Wahrscheinlichkeit*? Oder lässt sich der statistische Syllogismus in mein Modell der theoretischen Wahrscheinlichkeit integrieren, so dass letztlich *nur* ein Modell logischer Wahrscheinlichkeit übrig bleibt? Dies wäre natürlich erwünscht, und daher werde ich im Folgenden eine Vereinheitlichung versuchen, indem ich den statistischen Syllogismus innerhalb meines Ansatzes analysiere.

3. AUSSAGEN-LOGISCHE ANALYSE

Streng genommen besitzt der statistische Syllogismus zwar eine *quantoren-* bzw. *prädikatenlogische* Struktur, aber wichtig ist dabei auch die – einfachere – *aussagen-logische* bzw. *junktoren-logische* Struktur. Daher analysiere ich das Problem zunächst (quantitativ) aussagenlogisch. Und zwar liegt hier die Struktur des *Modus ponens* vor:

A impliziert B	$A \rightarrow B$	quantitativ:	$p(A \rightarrow B) = 1$
<u>A ist wahr</u>	<u>A</u>		$p(A) = 1$
B ist wahr	B		$p(B) = 1$

Der statistische Syllogismus müsste in der hier vorgestellten Symbolik – entsprechend einem *statistischen Modus ponens* – im *Beispiel* dann folgendermaßen lauten:

$$\begin{array}{l} p(A \rightarrow B) = 0,8 \\ \underline{p(A) = 1} \\ p(B) = 0,8 \end{array}$$

Anders geschrieben:

$$p(A \rightarrow B) = 0,8 \wedge p(A) = 1 \Rightarrow p(B) = 0,8$$

(Warum ich hier den Doppelpfeil für *vollständige* Deduktion nehme, wird gleich ersichtlich.)

Zur genaueren Erläuterung die Darstellung in Formeln, wie sie im Text entwickelt wurden:

$$\frac{a+c+d}{a+b+c+d} = 0,8 \wedge \frac{a+b}{a+b+c+d} = 1 \Rightarrow \frac{a+c}{a+b+c+d} = 0,8$$

Aus dem zweiten Ausdruck (der zweiten Prämisse) ergibt sich: $c = 0, d = 0$.

Somit müssen der erste und der dritte Ausdruck, also die erste Prämisse und die Konklusion, den gleichen Wert haben, nämlich 0,8.

• Diese Analyse zeigt zwei vielleicht überraschende Ergebnisse:

Erstens, $p(A \rightarrow B) = 0,8 \wedge p(A) = 1 \Rightarrow p(B) = 0,8$ ist kein „Wahrscheinlichkeits-Schluss“.

Hier liegt *kein partieller* Schluss vor, sondern ein *vollständiger* Schluss. Der Schluss-Satz ist vollständig aus den beiden Prämissen ableitbar, der Ableitungsgrad, d. h. die *theoretische* Wahrscheinlichkeit beträgt $p^T = 1$.

$$p^T[p(A \rightarrow B) = 0,8 \wedge p(A) = 1 \Rightarrow p(B) = 0,8] = 1 \quad \text{Anders dargestellt:}$$

$$\begin{array}{l} (A \rightarrow B) = 0,8 \\ p(A) = 1 \\ \hline p(B) = 0,8 \end{array} p^T = 1$$

Zweitens, nicht der Wert der *Ableitung* stimmt mit dem Wert der *ersten Prämisse* überein – die Ableitung besitzt eben die (*theoretische*) Wahrscheinlichkeit 1,0 ($p^T = 1,0$), sondern die *Konklusion* besitzt die gleiche (*empirische*) Wahrscheinlichkeit wie die erste Prämisse, nämlich 0,8, d. h. $p = 0,8$.

Nun sind A und B zunächst *unstrukturierte* Aussagen, es fragt sich, ob und wie weit sich solche Aussagen zur Analyse des statistischen Syllogismus verwenden lassen. Allerdings ist – wie beschrieben – nach meinem Verständnis die „Aussagen-Logik“ nicht auf Aussagen festgelegt, sondern A und B können auch für Individuen oder Klassen o. a. stehen. Das wäre dann für das Beispiel wie folgt darzustellen:

$$\begin{array}{l} (\text{Philosoph} \rightarrow \text{weise}) = 0,8 \\ p(\text{Philosoph}) = 1 \\ \hline p(\text{weise}) = 0,8 \end{array} p^T = 1$$

Hier könnte man einwenden, dass damit das Wesen des statistischen Syllogismus (bzw. des Beispiels) nicht erfasst wird, allein schon, weil eben gar kein *Eigennamen* oder keine Individuumskonstante auftaucht. Ich werde daher später den statistischen Syllogismus mittels *Prädikaten-Logik* analysieren, die auf *Individuen* Bezug nimmt.

Die Frage ist aber, ob in der *aussagen-logischen* Analyse nicht dennoch die wesentliche logische Struktur des statistischen Syllogismus erfasst wird. Allerdings geht es dabei aus der Sicht meines Modells um einen Sonderfall, wie ich im nächsten Punkt erläutere.

- Der *Modus ponens* mit $p(A) = 1$ ist ein *Sonderfall*.

Der *Modus ponens* wie im obigen Beispiel $p(A \rightarrow B) = 0,8 \wedge p(A) = 1 \Rightarrow p(B) = 0,8$ ist ein *Sonderfall*, weil hier gilt $p(A) = 1$, also die zweite Prämisse ist (positiv) *deterministisch*.

Unter dieser Bedingung gilt tatsächlich, dass man aus dem Wert der *ersten Prämisse* den Wert der *Konklusion* automatisch ablesen kann, wie wir es aus dem statistischen Syllogismus kennen, also $p(A \rightarrow B) = p(B)$.

$$\text{Allgemein gilt: } p^T[p(A \rightarrow B) = r/n \wedge p(A) = 1 \Rightarrow p(B) = r/n] = 1$$

Wenn $p(A) \neq 1$, ist die Berechnung jedoch wesentlich komplizierter. Ein wirklich statistischer Schluss mit der Struktur des Modus ponens, ein „echter“ *statistischer Syllogismus* könnte z. B. lauten:

$$p^T[p(A \rightarrow B) = 2/3 \wedge p(A) = 1/3 \longrightarrow p(B) = 1/3] = 58/64$$

Oder anders geschrieben:

$$\frac{p(A \rightarrow B) = 2/3}{p(A) = 1/3} \quad p^T = 58/64$$

$$p(B) = 1/3$$

In diesem Fall sind *alle* Werte *statistisch* (also $\neq 1$ und $\neq 0$): die Werte der beiden Prämissen, der Wert der Konklusion, und der Wert der Ableitung. Somit liegt auch nur ein *partieller* oder *semi-analytischer* oder eben *statistischer* Schluss vor.

Hier reicht es nicht, nur einen *Dezimalwert* wie z. B. 0,66 (statt 2/3) einzugeben, denn p^T hängt von der *absoluten* Größe ab, ist unterschiedlich, je nachdem, ob z. B. $p = 2/3, 4/6, 6/9$ usw. ist.

Nachfolgend erläutere ich eine Möglichkeit der Berechnung von p^T , ohne dass hier eine Formel entwickelt werden soll.

Zunächst ist aber zu klären, welche Werte $p(B)$ annehmen kann.

$$\frac{a+c+d}{a+b+c+d} = 2/3 \wedge \frac{a+b}{a+b+c+d} = 1/3 \longrightarrow \frac{a+c}{a+b+c+d} = ?$$

aus dem *ersten* Bruch: $a + c + d = 2, a + b + c + d = 3$, also $b = 1$
 aus dem *zweiten* Bruch: $a + b = 1, c + d = 2$
 zusammen bzw. Ergebnisse: $a = 0, b = 1, c + d = 2$

Das bedeutet: Für $p(B)$, als Bruch $\frac{a+c}{a+b+c+d}$, gibt es drei Möglichkeiten:

$c = 2, d = 0: p(B) = 2/3$
 $c = 1, d = 1: p(B) = 1/3$
 $c = 0, d = 2: p(B) = 0/3$

Zwar *könnte* in diesem Beispiel auch der Wert der Konklusion $p(B)$ mit dem Wert der ersten Prämisse $p(A \rightarrow B) = 2/3$ übereinstimmen. Aber dies ist nicht notwendig (und wäre zufällig).

Und in keinem Fall stimmen der Wert der ersten Prämisse $p(A \rightarrow B) = 2/3$ und der Wert der Ableitung p^T überein.

Wählen wir aber die Möglichkeit $p(B) = 1/3$ und bestimmen den Grad der Ableitung: p^T .

Die Frage ist also: wenn $p(A \rightarrow B) = 2/3$ und $p(A) = 1/3$, wie wahrscheinlich ist es dann, dass $p(B) = 1/3$?

Die Werte von a ($a = 0$) und b ($b = 1$) stehen fest, aber wir müssen zwischen c und d differenzieren, wobei 3 Elemente x_1, x_2 und x_3 zu unterscheiden sind, da $a + b + c + d = n = 3$.

Nr.	a	b	c	d
1	0	x_3	x_1, x_2	0
2	0	x_2	x_1, x_3	0
3	0	x_1	x_2, x_3	0
4	0	x_3	x_1	x_2
5	0	x_3	x_2	x_1
6	0	x_2	x_1	x_3
7	0	x_2	x_3	x_1
8	0	x_1	x_2	x_3
9	0	x_1	x_3	x_2
10	0	x_3	0	x_1, x_2
11	0	x_2	0	x_1, x_3
12	0	x_1	0	x_2, x_3

Zunächst die einfachere Berechnung für die *Positiv-Implikation* $* \rightarrow$:

Es gibt 12 Möglichkeiten, dass die Prämissen wahr sind, dass $a = 0$, $b = 1$ und $c + d = 2$.

Wir müssen klären, in wie vielen von diesen 12 Fällen gilt: $c = 1$

In den *Zeilen* 4 bis 9 steht 1 x unter c , d. h. von 12 Möglichkeiten ist in 6 Fällen $c = 1$.
Somit $p^T = 6/12$.

Man kann also im Ganzen formulieren:

$$p^T[p(A \rightarrow B) = 2/3 \wedge p(A) = 1/3 * \rightarrow p(B) = 1/3] = 6/12 = 1/2 = 0,5$$

Nun zur Berechnung für die *Normal-Implikation* \rightarrow :

Eine Implikation ist nur falsch, wenn das *Vorderglied wahr* ist und das *Nachglied falsch*.

D. h. hier: $p(A \rightarrow B) = 2/3 \wedge p(A) = 1/3 \rightarrow p(B) = 1/3$ ist nur falsch, wenn

$p(A \rightarrow B) = 2/3 \wedge p(A) = 1/3$ wahr ist und $p(B) = 1/3$ falsch ist (also $p(B) = 2/3$ oder $0/3$).

Das ist in den Zeilen 1 bis 3 und 9 bis 12 der Fall, also in 6 Möglichkeiten.

Die Gesamtanzahl ist aber hier nicht 12, sondern 64 (4^3), weil eben nicht nur die Fälle zählen, in denen die Prämissen *wahr* sind (wie bei der Positiv-Implikation), sondern alle Möglichkeiten.

Die 6 *negativen* Fälle zieht man ab von *allen* Fällen, also von 64.

So erhält man: $p^T = 64/64 - 6/64 = 58/64 = 0,91$

Man kann somit vollständig formulieren:

$$p^T[p(A \rightarrow B) = 2/3 \wedge p(A) = 1/3 \rightarrow p(B) = 1/3] = 58/64 = 29/32 = 0,91$$

4. PRÄDIKATEN-LOGISCHE ANALYSE MITTELS IMPLIKATION

Kehren wir jetzt zurück zum statistischen Syllogismus im üblichen Sinne, bei dem wie beschrieben gilt:

- Die erste Prämisse ist statistisch: $0 < p < 1$
- Die zweite Prämisse ist deterministisch: $p = 1$.
- Der Schluss-Satz ist statistisch, und sein Wert stimmt überein mit der ersten Prämisse.

Genauer kann man den statistischen Syllogismus darstellen, wenn man ihn *prädikatenlogisch* analysiert. Denn er ist eine logische Form, bei der (anders als in der Aussagen-Logik) die Innen-Struktur der Sätze, das Verhältnis von Subjekt und Prädikat, eine Rolle spielt.

Nehmen wir einen Satz wie z. B. ‚Sokrates ist weise‘. Ein solcher Satz hat eine prädikatenlogische Struktur, weil hier einem Subjekt (Argument) ein Prädikat zugeordnet wird. Man schreibt normalerweise logisch: Fx_i oder $x_i \in F$. Das Problem ist, dass sich, bei dieser Formalisierung, prädikatenlogische Strukturen nicht ohne weiteres in mathematische Formeln übersetzen lassen.

Es bieten sich daher zwei alternative Formalisierungen an:

1. mittels der *Implikation*: ‚Sokrates ist weise‘ schreibt man dann: Sokrates \rightarrow weise
Formal $x_i \rightarrow F$ bzw. quantitativ $p(x_i \rightarrow F) = 1$
2. mittels der *Positiv-Implikation*: Sokrates $*\rightarrow$ weise
Formal: $x_i *\rightarrow F$ bzw. quantitativ $p(x_i *\rightarrow F) = 1$

In diesem 4. Punkt verwende ich erst einmal die übliche Implikation \rightarrow :

Es stellt sich also die Frage, ob die folgende *Hypothese* gültig ist. Gilt

$$p(F \rightarrow G) = r/n \wedge p(x_i \rightarrow F) = 1 \Rightarrow p(x_i \rightarrow G) = r/n ?$$

Ich übersetze diesen Ausdruck wieder in mathematische Formeln. Da wir es hier mit 3 Ausdrücken F , G und x zu tun haben, ergeben sich *umfangreichere Formeln* (mit 8 Zahl-Variablen), die sich aus den Wahrheitstafeln ableiten lassen.

Die entsprechenden Wahrheitstafeln sind:

<u>F</u>	<u>G</u>	<u>x_i</u>		$F \rightarrow G$	$x_i \rightarrow F$	$x_i \rightarrow G$
+	+	+	a	+	+	+
+	+	-	b	+	+	+
+	-	+	c	-	+	-
+	-	-	d	-	+	+
-	+	+	e	+	-	+
-	+	-	f	+	+	+
-	-	+	g	+	-	-
-	-	-	h	+	+	+

Wir erhalten also 3 Formeln:

$$(1) \quad p(F \rightarrow G) = \frac{a + b + e + f + g + h}{a + b + c + d + e + f + g + h} = \frac{r}{n}$$

$$(2) \quad p(x_i \rightarrow F) = \frac{a + b + c + d + f + h}{a + b + c + d + e + f + g + h} = 1$$

$$(3) p(x_i \rightarrow G) = \frac{a+b+d+e+f+h}{a+b+c+d+e+f+g+h} = \frac{s}{n}$$

1. Fall: $p(F \rightarrow G) = 1$

Prüfen wir die Gleichungen zunächst für den einfachen, nämlich *deterministischen* Syllogismus, bei dem auch die erste Prämisse $p(F \rightarrow G)$ den Wert $p = 1$ hat.

- Dann ergibt sich aus der (1) Gleichung: $c + d = 0$
- Und aus der (2) Gleichung: $e + g = 0$
- Dann bleibt bei der (3) Gleichung stehen:

$$\frac{a+b+f+h}{a+b+f+h} = 1 \text{ Dies ist also das erwünschte Resultat. Denn hier gilt:}$$

$$p(F \rightarrow G) = r/n = 1 \wedge p(x_i \rightarrow F) = 1 \Rightarrow p(x_i \rightarrow G) = r/n = 1$$

Es ist also ein *strenger* Schluss gegeben, mit $p^T = 1$.

2. Fall: $p(F \rightarrow G) = r/n$, dabei gilt: $0 < p < 1$

Hier ist $p(F \rightarrow G)$ also variabel bzw. statistisch. Sonst gilt wieder $p(x_i \rightarrow F) = 1$.

Gemäß dem statistischen Syllogismus müsste man erwarten, dass dann gilt: $p(Fx_i \rightarrow G) = r/n$.

Das lässt sich aber nicht aus den Formeln ableiten.

- Aus der (2) Formel ergibt sich wie gesagt: $e = 0, g = 0$
- Aus der (1) Formel ergibt sich: $a + b + e + f + g + h = r$.
Da $e = 0, g = 0$, folgt daraus: $a + b + f + h = r$
- Bei der (3) Formel ist der Zähler: $a + b + d + e + f + h$. Da $e = 0$, bleibt: $a + b + d + f + h$.
Wenn gelten sollte $a + b + d + f + h = r$, dann müsste $d = 0$ sein. Das steht aber nicht fest.
Das Problem ist der *unbestimmte* Wert von d . Man kann nur sagen: $c + d = n - r$. Somit
 $d = n - r - c$

$p(F \rightarrow G) = r/n \wedge p(x_i \rightarrow F) = 1 \longrightarrow p(x_i \rightarrow G) = r/n$ ist also nur ein *semi-analytischer* Schluss ($0 < p^T < 1$), kein strenger Schluss.

3. Fall: $p(F \rightarrow G) = 4/5$

Versuch, ob sich durch Einsetzung *konkreter* (statistischer) Zahlen für $p(F \rightarrow G)$ eine Lösung ergibt.

Aus obiger Gleichung ergibt sich: $c + d = 1$. Also kann man den Wert von d auch hier nicht genau bestimmen. Es sind möglich: $d = 1$ oder $d = 0$. Nur wenn $d = 0$, hat $p(x_i \rightarrow G)$ den gleichen Wert wie $p(F \rightarrow G)$, nämlich $4/5$. Nur dann liegt ein typischer statistischer Syllogismus vor. Wenn $d = 1$, ergibt sich dagegen $p(x_i \rightarrow G) = 5/5$.

Es ergibt sich (bei $d = 0$):

$$\begin{array}{r} p(F \rightarrow G) = 4/5 \\ p(x_i \rightarrow F) = 1 \\ \hline p(x_i \rightarrow G) = 4/5 \end{array} p^T < 1$$

D. h. auch hier gibt es keinen deterministischen, keinen vollständiger Schluss. (Wie die logische Wahrscheinlichkeit p^T oben genau zu berechnen ist, soll nicht erläutert werden.)

4. Zusätzliche Prämisse

Man könnte postulieren, dass noch eine Prämisse $p(x_i) = 1$ einzuführen ist. Wie unten gezeigt wird, ergibt sich in diesem Fall in der Tat ein *deterministischer* Schluss, und der Schluss-Satz hat den Wert $p = r/n$ bzw. im Beispiel $4/5$, wie die erste Prämisse. Das entspricht also genau dem Modell des statistischen Syllogismus.

Zur Erklärung:

$$p(x_i) = 1 \text{ bedeutet: } \frac{a + c + e + g}{a + b + c + d + e + f + g + h} = 1$$

Daraus resultiert: $b = 0, d = 0, f = 0, h = 0$. Aus der (2) Formel ergab sich: $e = 0, g = 0$.

Somit folgt für die Formel (3): $\frac{a}{a + c} = r/n$

Und exakt dies ergibt sich auch für die Formel (1)

$$p(F \rightarrow G) = \frac{a + b + e + f + g + h}{a + b + c + d + e + f + g + h} = \frac{r}{n}$$

wenn man alle Variablen mit dem Wert 0 herausstreicht. Man erhält also:

$$\begin{array}{l} p(F \rightarrow G) = r/n \\ p(x_i \rightarrow F) = 1 \\ p(x_i) = 1 \\ \hline p(x_i \rightarrow G) = r/n \end{array} p^T = 1$$

5. PRÄDIKATEN-LOGISCHE ANALYSE MITTELS POSITIV-IMPLIKATION

Da die Implikation ja wie beschrieben einige Probleme aufwirft, ist zu prüfen, wie sich der statistische Syllogismus mittels der *Positiv-Implikation* $*\rightarrow$ darstellen lässt.

Gilt hier die Hypothese:

$$p(F *\rightarrow G) = r/n \wedge p(x_i *\rightarrow F) = 1 \quad *\Rightarrow \quad p(x_i *\rightarrow G) = r/n ?$$

Ich übersetze diesen Ausdruck wieder in *mathematische* Formeln, entsprechend den Formeln für die Positiv-Implikation, wie sie im Text aus den Wahrheitstafeln entwickelt wurden.

$$(1) \quad p(F \rightarrow G) = \frac{a + b}{a + b + c + d} = \frac{r}{n}$$

$$(2) p(x_i \rightarrow F) = \frac{a+c}{a+c+e+g} = 1$$

$$(3) p(x_i \rightarrow G) = \frac{a+e}{a+c+e+g} = \frac{s}{m}$$

Unterscheiden wir wieder zwei Fälle:

1. Fall: $p(F \ast \rightarrow G) = 1$

Prüfen wir die Gleichungen zunächst wieder für den einfachsten, *deterministischen* Fall, dass auch der Wert für $p(F \ast \rightarrow G) = 1$ ist.

- Dann ergibt sich aus der (1) Gleichung: $c + d = 0$
- Aus der (2) Gleichung: $e + g = 0$
- Somit ergibt sich für die (3) Formel:

$$\frac{a+0}{a+0+0+0} = \frac{a}{a} = 1$$

Für den deterministischen Fall stimmt also die Formel

$$p(F \ast \rightarrow G) = 1 \wedge p(x_i \ast \rightarrow F) = 1 \ast \Rightarrow p(x \ast \rightarrow G) = 1. \text{ Entsprechend } p^T = 1.$$

2. Fall: $p(F \ast \rightarrow G) = r/n$, dabei $0 < p < 1$

Hier ist $p(F \ast \rightarrow G)$ also variabel bzw. statistisch.

Gemäß dem statistischen Syllogismus müsste man erwarten, dass dann gilt: $p(x_i \rightarrow G) = r/n$. Bei der normalen Implikation lässt sich das aber nicht aus den Formeln ableiten.

Wie sieht es bei der *Positiv-Implikation* aus?

- Aus der (1) Gleichung ergibt sich z. B.: $a + b = r$, also $a = r - b$
- Aus der (2) Gleichung ergibt sich wie gesagt: $e = 0$, $g = 0$

- Daraus folgt für die Gleichung (3): $\frac{a+e}{a+c+e+g} = \frac{s}{m}$

$$\frac{a}{a+c} = \frac{s}{m} \quad \text{Vergleichen wir die mit der (1) Gleichung } \frac{a+b}{a+b+c+d} = \frac{r}{n}$$

Damit ist aber die Übereinstimmung von (1) und (3) keineswegs gesichert. Der *Zähler* wäre nur gleich, wenn $b = 0$. Der *Nenner* wäre nur gleich, wenn zusätzlich $d = 0$; beides ist nicht beweisbar.

D. h. auch mit der *Positiv-Implikation* gilt die folgende Ableitung nur als *partieller Schluss*:

$$p(F \ast \rightarrow G) = r/n \wedge p(x_i \ast \rightarrow F) = 1 \ast \longrightarrow p(x_i \ast \rightarrow G) = r/n$$

Denn nur wenn $b = 0 \wedge d = 0$, ist der Wert von Gleichung (1) und der Wert von Gleichung (3) gleich.

Auch hier gilt die Ableitung nur, wenn man eine *Zusatzhypothese* $p(x_i) = 1$ einführt, also:

$$p(F \ast \rightarrow G) = r/n \wedge p(x_i \ast \rightarrow F) = 1 \wedge p(x_i) = 1 \ast \Rightarrow p(x_i \ast \rightarrow G) = r/n$$

Diese prädikaten-logische Analyse – mit Zusatzhypothese – führt somit wie gewünscht zu einem gültigen Schluss, bei dem gilt: $p(F \rightarrow G) = p(x_i \rightarrow G)$, entsprechend auch für die Positiv-Implikation.

BILANZ: STATISTISCHER SYLLOGISMUS

1) ERGEBNISSE

Als *statistischen Syllogismus* kann man eine Schluss-Form bezeichnen, die – nach meiner Analyse – folgende Eigenschaften besitzt:

- Die erste Prämisse ist statistisch: $p = r/n$, $0 < p < 1$.
- Die zweite Prämisse ist deterministisch: $p = 1$.
- Der Schluss-Satz ist statistisch, und sein Wert stimmt überein mit der ersten Prämisse.
- Die Ableitung selbst, mittels *theoretischer Wahrscheinlichkeit*, ist deterministisch: $p^T = 1$

Wichtig ist: Anders als sonst dargestellt, besitzt der Schluss-Satz eine *empirische* (statistische) Wahrscheinlichkeit $0 < p < 1$, *keine logische*. Dagegen ist die Deduktion *deterministisch*, sie besitzt eine *theoretische* (oder *logische*) Wahrscheinlichkeit $p^T = 1$.

2) AUSSAGEN-LOGISCHE ANALYSE

Aussagen-logisch lässt sich die Formalisierung des statistischen Syllogismus überzeugend durchführen:

$$p^T[p(A \rightarrow B) = r/n \wedge p(A) = 1 \Rightarrow p(B) = r/n] = 1$$

Aber diese Formalisierung berücksichtigt nicht konkret *individuelle* Sätze, wie sie im statistischen Syllogismus vorkommen.

3) PRÄDIKATEN-LOGISCHE ANALYSE

Prädikaten-logisch werden die individuellen Sätze erfasst, aber man erhält zunächst nur einen *partiellen* Schluss:

$$p(F \rightarrow G) = r/n \wedge p(x_i \rightarrow F) = 1 \longrightarrow p(x_i \rightarrow G) = r/n, \text{ also } 0 < p^T < 1$$

Doch durch Einführung einer *Zusatz-Prämisse* $p(x_i) = 1$ wird der Schluss *streng gültig*:

$$p(F \rightarrow G) = r/n \wedge p(x_i \rightarrow F) = 1 \wedge p(x_i) = 1 \Rightarrow p(x_i \rightarrow G) = r/n, \text{ also } p^T = 1$$

Bei Verwendung der *Positiv-Implikation* $*\rightarrow$ wird die Situation nicht verändert.

Es bleiben allerdings einige *Probleme* bestehen:

- Die Formalisierung eines individuellen Satzes wie ‚Sokrates ist Philosoph‘ mittels des aussagen-logischen *Implikators* als $x_i \rightarrow F$.

Im Punkt 0-4 (vor allem 0-4-3-2) habe ich aber erläutert, warum dies m. E. legitim ist.

- Die *quantitative Darstellung* eines solchen individuellen Satzes mit der empirischen Wahrscheinlichkeit p als $p(x_i * \rightarrow G) = 1$

Wie ist ein solcher Satz zu interpretieren? Man könnte ihn z. B. verstehen als ‚Sokrates ist in allen Fällen Philosoph‘ (vgl. hierzu 1-4-5, vor allem 1-4-5-1 und unten im Text).

- Die Zuordnung eines Wahrheitswertes für die *Zusatz-Hypothese* ‚ x_i ‘.

Man könnte einwenden, ‚ x_i ‘ ist kein Satz, sondern ein Zeichen, man kann ihm keinen Wahrheitswert zuordnen. Wie ich aber in 0-4-4 gezeigt habe, ist das im Rahmen einer *funktionalen* Logik doch möglich. ‚ x_i ‘ steht dann für ‚ x_i existiert‘ oder ‚ x_i ist belegt‘.

- Die *quantitative Darstellung* der Zusatz-Hypothese als $p(x_i) = 1$.

Wie ist das zu verstehen? Man könnte es im Beispiel so deuten, als ‚Sokrates existiert in allen Fällen‘.

4) DIE ZUORDNUNG DES STATISTISCHEN WERTES

Es geht hier genau um die Zuordnung der *statistischen* Wahrscheinlichkeit zum *Schluss-Satz*, *nicht zur Ableitung*.

Die Analyse dieses Problem ist sehr komplex und richtet sich vor allem an Experten. Kehren wir zur Einstieg noch einmal zum Anfangsbeispiel zurück:

80% der F sind G	80% der Philosophen sind weise
<u>x_i ist ein F</u>	<u>Sokrates ist ein Philosoph</u>
x_i ist mit $p = 0,8$ ein G	Mit 80% Wahrscheinlichkeit ist Sokrates weise

Dies wäre nach dem *herkömmlichen* Verständnis des statistischen Syllogismus folgendermaßen zu deuten: Die Wahrscheinlichkeit (z. B. 0,8) – als *logische* Wahrscheinlichkeit – wird der *Ableitung* zugeschrieben, nicht dem *Schluss-Satz* ‚ x_i ist ein G‘.

Dieser individuelle Schluss-Satz ‚ x_i ist ein G‘ wird *gar nicht quantifiziert*, genauso wie die individuelle Prämisse ‚ x_i ist ein F‘.

Bei allen Formalisierungen ergab sich jedoch das Umgekehrte:

Der Schluss-Satz ist vollständig aus den beiden Prämissen ableitbar, der Ableitungsgrad, d. h. die *theoretische* oder *logische* Wahrscheinlichkeit beträgt $p^T = 1$.

Und nicht der Wert der *Ableitung* stimmt mit dem Wert der *ersten Prämisse* überein, sondern die *Konklusion* besitzt die gleiche (*empirische*) Wahrscheinlichkeit wie die erste Prämisse, nämlich 0,8, d. h. $p = 0,8$.

Wenn dem aber so ist, wie lässt sich ein *individueller* Satz wie ‚ x_i ist mit $p = 0,8$ ein F‘ oder ‚Sokrates ist mit 80% Wahrscheinlichkeit weise‘ deuten?

- *Variabler* individueller Satz

Nehmen wir zunächst einen *variablen* individuellen Satz wie: ‚Ein beliebiger (oder jeder beliebige) Philosoph ist mit 80% Wahrscheinlichkeit weise‘. Und setzen ihn in Relation zu einem allgemeinen statistischen Satz:

‚80% der Philosophen sind weise‘.

‚Ein *beliebiger* Philosoph ist mit 80% Wahrscheinlichkeit weise‘.

Man könnte denken, der zweite Satz ist ein *Schluss* aus dem ersten Satz, aber dem ist nicht so. Die beiden Sätze sind einfach *synonyme* Umformungen oder Umformulierungen (man könnte auch sagen Äquivalenzen).

Der zweite Satz ist also *keine* Ableitung aus dem ersten mit $p^T = 80\%$, sondern auch hier gilt in beiden Fällen: $p = 0,8$, *nicht* $p^T = 0,8$. Nur ist der erste Satz mit der *relativen Häufigkeit* formuliert, der zweite mit der entsprechenden *Wahrscheinlichkeits*-Angabe.

- *Konstanter* individueller Satz

Nun zu einem *bestimmten* individuellen Satz, wie er im statistischen Syllogismus verwendet wird, z. B. ‚Sokrates ist mit 80% Wahrscheinlichkeit weise‘. Welche Interpretationen sind hier möglich?

– *Umformung*

Ist hier auch denkbar, dass der Satz nur eine Umformung ist? Wohl kaum, $p(x_i \rightarrow G) = 0,8$ ist nicht äquivalent $p(F \rightarrow G) = 0,8 \wedge p(x_i \rightarrow F) = 1$ (bzw. mit Zusatzhypothese $p(x_i) = 1$).

Allerdings, wenn man die singuläre Prämisse dem Schluss-Satz hinzufügt, gilt doch generell

$A \rightarrow B \wedge A \Leftrightarrow A \wedge B$ bzw. $p(A \rightarrow B) = r/n \wedge p(A) = 1 \Leftrightarrow p(A) = 1 \wedge p(B) = r/n$, aber das will ich nicht weiter verfolgen.

– *Umdeutung*

Eine weitere Frage ist: Wenn gilt: $p(x_i \rightarrow G) = 0,8$ folgt mit $p^T = 1$ aus den Prämissen, könnte man das umdeuten in: $p(x_i \rightarrow G) = 1$ folgt mit $p^T = 0,8$ aus den Prämissen? Dann wäre also der Schluss-Satz *deterministisch* und die Ableitung *statistisch* – und das scheint unproblematisch.

tischer zu handhaben. Aber diese Umformung ist falsch, wenn $p(x_i \rightarrow G) = 0,8$ *sicher* folgt ($p^T = 1$), dann folgt $p(x_i \rightarrow G) = 1$ nur mit einer Wahrscheinlichkeit $p^T = 0$. Außerdem, das Grundproblem, wie ein quantitativer (bestimmter) individueller Satz zu deuten ist, ergibt sich vergleichbar bei $p = 1$ wie bei z. B. $p = 0,8$.

Allenfalls, wenn man $p(x_i \rightarrow G) = 1$ quasi *de-quantifizieren* würde, zu „ $x_i \rightarrow G$ “; entsprechend $p(x_i \rightarrow G) = 0$ zu $\neg(x_i \rightarrow G)$, entsprechend der qualitativen Aussagen-Logik. Dann wäre das Problem der Quantifizierung individueller Sätze erledigt. Nur dann könnte man eben keine *mathematisch-quantitative* Berechnung des statistischen Syllogismus mehr vornehmen, wie sie hier gerade angestrebt wird.

– quantitative Wahrheit

Könnte man ‚ $p(x_i \rightarrow G) = 0,8$ ‘ deuten als: ‚Der Satz ‚ $x_i \rightarrow G$ ‘ ist zu 80% (empirisch) wahr, hat einen *Wahrheitsgrad* von 0,8‘? Normalerweise wird die (empirische) Wahrheit ja nicht quantifiziert, man unterscheidet nur *qualitativ* zwischen *wahr* und *falsch*. Ich habe aber gezeigt, dass es unter bestimmten Bedingungen durchaus sinnvoll ist, Wahrheit zu quantifizieren (vgl. 1-3-0-2). Das wäre hier wie folgt zu deuten:

Der reale Sachverhalt sei: $p(x_i \rightarrow G) = 1,0$.

Wenn der Satz aussagt, $p(x_i \rightarrow G) = 0,8$, dann hat er einen Wahrheitsgrad von 0,8.

Diese Argumentation ist sinnvoll, aber sie trifft nicht das, was mit dem statistischen Syllogismus intendiert ist; dort ist nicht gemeint, dass der Schluss-Satz einen bestimmten Wahrheits-Grad besitzt.

– intensionale Quantität

In eine ähnliche Richtung geht der Ansatz der intensionalen Quantität (vgl. vor allem 1-4-5). Man deutet p hier so, dass damit etwas über das *Ausmaß der Eigenschaft* ausgesagt wird. Am konkreten Beispiel: Man würde dann nicht sagen ‚Sokrates ist *mit* 80% (Wahrscheinlichkeit) weise‘, sondern ‚Sokrates ist *zu* 80% weise‘. Aber auch das trifft nicht die Intentionen des statistischen Syllogismus.

– extensionale Quantität

Man interpretiert hier ‚ $p(x_i \rightarrow G) = 0,8$ ‘, indem man z. B. über die *Zeit* quantifiziert. Im Beispiel hieße das: ‚Sokrates ist in 80% seiner Lebenszeit weise, in 20% aber nicht‘. Das mag bei einer – eher konstanten – Eigenschaft wie Weisheit nicht sehr plausibel sein, aber bei einer anderen Eigenschaft wie z. B. ‚gut gelaunt‘ durchaus vorstellbar. In der Tat interpretiert man quantitative Individual-Sätze in dieser Art am besten, mit Bezug auf Zeit. Ähnlich könnte man sagen: In 80% aller *Fälle* ist Sokrates gut gelaunt. Oder unter 80% aller *Bedingungen* (vgl. 1-4-5-1).

Man mag zunächst denken, dass man hier auch von *möglichen Welten* sprechen könnte. ‚Sokrates ist mit 80% Wahrscheinlichkeit weise‘ wäre dann wie folgt zu verstehen: ‚Sokrates ist in 80% der möglichen Welten weise, in 20 % aber nicht‘. Aber Vorsicht, die Rede von möglichen Welten bezieht sich auf die *theoretische* Wahrscheinlichkeit (oder theoretische Wahrheit), nicht auf die empirische.

Jedenfalls ist die extensionale Quantität die angemessene Form, um einen *unabhängigen*, quantitativen, singulären Satz darzustellen. Man kann hier von *objektiver* oder *absoluter* Wahrscheinlichkeit sprechen (was nichts daran ändert, dass Wahrscheinlichkeit eine relative Größe ist).

Allerdings ist auch dieser Ansatz dem statistischen Syllogismus nicht angemessen. Der Schluss-Satz im Beispiel ist keineswegs eine Aussage über Sokrates Verhalten zu unterschiedlichen Zeiten, sondern eine *Ableitung von den Häufigkeitsverhältnissen* bei allen Philosophen, von denen eben 80% weise sein sollen.

– *subjektive Wahrscheinlichkeit*

Am richtigsten ist es daher, die Wahrscheinlichkeit bei der Konklusion zu *relativieren*.

Eine Möglichkeit ist, von einer *subjektiven* Wahrscheinlichkeit auszugehen. Man argumentiert: real, *objektiv*, gilt *entweder* Sokrates ist weise *oder* er ist nicht weise (d. h. quantitativ: er ist mit 100% Wahrscheinlichkeit weise oder er ist mit 0% Wahrscheinlichkeit weise). Wir wissen aber *subjektiv* nicht sicher, welche der beiden Möglichkeiten zutrifft, wir wissen nur mit 80% Sicherheit, dass er weise ist. Wenn wir allein wissen, dass Sokrates Philosoph ist, aber sonst keine weitere, spezifische Information über ihn haben, dann sind die 80% eben die einzige Aussage, die wir treffen können.

– *bedingte Wahrscheinlichkeit*

Am besten ist es aber, beim statistischen Syllogismus von einer *relativen* oder *bedingten* Wahrscheinlichkeit auszugehen.

D. h. man kann einen Satz wie ‚Sokrates ist mit 80% Wahrscheinlichkeit weise‘ gar nicht für sich *allein* deuten, sondern nur *in Relation* zu den Prämissen.

Die 80% im Schluss-Satz sagen nichts über das *spezifische* Individuum Sokrates aus, sondern, wie die Oberprämisse, letztlich nur etwas über die *Klasse* der Philosophen. Für einen (jeden) anderen Philosophen, z. B. Platon, ergäben sich die gleichen 80% Wahrscheinlichkeit. Es wird eigentlich nur ausgesagt, dass ein Individuum zu einer bestimmten Gruppe (Klasse) gehört, für die eine bestimmte *Häufigkeitsverteilung* gilt. Nur wenn für diese Gruppe 100% oder 0% gelten würden, dann wäre auch das Individuum *eindeutig* bestimmt.

Diese bedingte Wahrscheinlichkeit könnte man wie folgt schreiben:

$$p(\text{Sokrates ist weise, 80\% der Philosophen sind weise} \wedge \text{Sokrates ist Philosoph}) = 0,8.$$

Aber wie erläutert, es ist die *empirische Wahrscheinlichkeit* p , mit p^T könnte man schreiben.

$$p^T[p(\text{Sokrates ist weise, 80\% der Philosophen sind weise} \wedge \text{Sokrates ist Philosoph}) = 0,8] = 1$$

Hier sind noch Fragen bzw. Antworten offen. Dennoch glaube ich sagen zu können: Es ist gelungen, ein mathematisches Modell des statistischen Syllogismus anzugeben. Und es ist gelungen, den statistischen Syllogismus in meine Theorie zu integrieren, nach welcher der *Grad der Folgerichtigkeit* mittels der *theoretischen Wahrscheinlichkeit* p^T berechnet wird.