RUDOLF CARNAP

NOTES ON PROBABILITY AND INDUCTION*

INTRODUCTORY REMARKS

I. *The Three Main Conceptions of Probability*

(1) The classical conception (Bernoulli, Bayes, Laplace).
(2) The frequency conception (Mises, Reichenbach; mathematical statistics).
(3). The logical conception (Keynes, Jeffreys).
Read: Nagel [37]

II. *The Two Explicanda*

There are two explicanda, both called ‘probability’:
(1) logical or inductive probability (probability$_1$),
(2). statistical probability (probability$_2$).
Read: [Prob.] Ch. II, esp. §§ 9 and 10.
The logical concept of probability appears in three forms ([Prob.] § 8):
(a) the classificatory concept (confirming evidence),
(b) the comparative concept (higher confirmation),
(c) the quantitative concept (degree of confirmation).

III. *Preliminary Remarks on Inductive Logic*

Read: [Prob.] Ch. IV. In particular:
(1) Logical probability (as explicandum) is explained as a *fair betting quotient*, and as an estimate of relative frequency ([Prob.] § 41).
(2) If logical probability is used, no synthetic assumption (e.g., *uniformity* of the world) is needed as presupposition for the validity of the inductive method ([Prob.] § 41 F).
(3) Comparison of *inductive and deductive logic* ([Prob.] § 43).
(4) The main kinds of *inductive inference* ([Prob.] § 44 B)
(a) direct inference (from the population to a sample),

* University of California, Los .Angeles, Philosophy 242, Fall Semester 1955. Edited for publication by Arthur Benson.

Synthese 25 (1973) 269-298.

(b) predictive inference (from one sample to another),
(c) inference by analogy,
(d) inverse inference (from a sample to the population),
(e) universal inference (from a sample to a universal law).

(5) The use of inductive logic for the choice of a *practical decision* ([Prob.] §§ 50, 51). The rule of maximizing the estimate of utility. Daniel Bernoulli's law of utility.

IV. *Some Concepts of Deductive Logic*

Read: [Prob.] §§ 14-20, esp. 18-20.

(1) *State-descriptions* (*Z*, § 18A; comp. individual distributions, D 26-6a). A state-description describes a (possible) state or *model*.

(2) The requirement of the logical independence of the primitive predicates (§ 18B) can be abandoned if the dependences are expressed by *meaning postulates* (see [15].)

(3) *Families of related primitive predicates* (§ 18C).

(4) The *range* of a sentence (§§ 18D, 19).

(5). *L-concepts* (§ 20). I write '$\vdash i$' for '*i* is L-true', and hence '$\vdash i \supset j$' for '*i* L-implies *j*', and '$\vdash \mathrm{i} \equiv \mathrm{j}$' for '*i* is L-equivalent to *j*'.

In simple languages (e.g., those used in [Prob.]) every model is describable by a state-description. In richer languages this is not possible; here the definitions of L-concepts and degree of confirmation are to be based on *models* rather than state-descriptions.

THE THEORY OF DEGREE OF CONFIRMATION

V. *Fundamental axioms* (Al-A5)

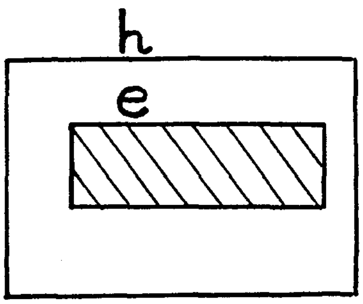
The axioms apply to any sentences *e* and *h* in a given language L (finite; or infinite). We presuppose throughout that the second argument of *c* (usually *e*) is not L-false (see [Prob.] pp. 295f.).

Al. *Range of values*. $0 \leq c\,(h, e) \leq 1$.
A2. *L-implication*. If $\vdash e \supset h$, then $c\,(h, e) = 1$.
A3. *Specialaddition principle*. If $e \,.\, h \,.\, h'$ is L-false, then $c\,(h \vee h', e) = c\,(h, e) + c\,(h', e)$.
A4. *General multiplication principle*. $c\,(h, h', e) = c\,(h, e) \times c\,(h', e \,.\, h)$.

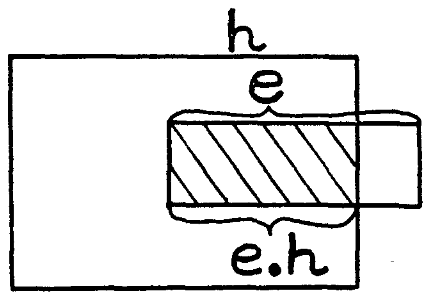
A5. *L-equivalent arguments*. If $\vdash e \equiv e'$ and $\vdash h \equiv h'$, then $c\,(h, e) = c\,(h', e')$.

(These axioms, except for A5, are those of Shimony [47]. They are together equivalent to the Conventions C53-1 and 2 in [Prob.] § 53. Most axiom systems of other authors are essentially equivalent to this one; see [Prob.] § 62.) The usual theorems of the probability calculus are provable on the basis of these axioms. Among them are the theorems [Prob.] T53-1a to f. ('*t*' is the tautology.)

VI. *Regular m-Functions and c-Functions* (A6)



Deductive logic
'*e* L-implies *h*' means that the range of *e* is entirely contained in that of *h*.



Inductive logic
'$c\,(h, e) = 3/4$' means that three-fourths of the range of *e* is contained in that of *h*. ([Prob.] § 55B.)

Fig. 1.

For degrees of confirmation (d. of c.) we need a measure function for the ranges of sentences. For this purpose we define regular *m*-functions ([Prob.] 55A).

Dl. *m* is a regular *m*-function for $\mathrm{L}_N =_{\mathrm{Df}}$
(a) for every Z_i in L_N, $m\,(Z_i) > 0$;
(b) $\sum_i m\,(Z_i) = 1$;
(c) if *j* is L-false, $m\,(j) = 0$;
(d) if *j* is not L-false, $m\,(j) = \sum m\,(Z_i)$ for all Z_i in the range of *j*.

D2. *c* is a regular *c*-function for $\mathrm{L}_N =_{\mathrm{Df}}$ there is a regular *m*-function *m* such that *c* is based upon *m*, i.e.

$$c\,(h, e) = \frac{m(e.h)}{m(e)}.$$

A6. *Regularity*. In a finite domain of individuals, $c(h, e) = 1$ *only* if $\vdash e \supset h$.

(This axiom corresponds to [Prob.] C53-3.)
Null confirmation is the d. of c. on the tautological evidence t ([Prob.] D57-1, where the symbol 'c_0' is used):

D3. $c_t(j) =_{\text{Df}} c(j, t)$

T1. A c-function c for L_N satisfies the axioms Al-A6 if and only if c is a regular c-function.

Proof. 1. Let c be a regular c-function for L_N. Then c satisfies Al-A6 according to [Prob.] T59-1a, 1b, 1l, 1n, 1h and i, T59-5a, respectively. 2. Let c satisfy Al-A6. Then c_t is a regular m-function (by [Prob.] C53-3 and T53-1). c is based upon c_t (comp. [Prob.] § 54B, (3)). Therefore c is a regular c-function.

According to Tl, the theorems stated in [Prob.] §§ 55, 57A and B, 59, 60, and 61 for regular c-functions in finite systems L_N are provable on the basis of Al to A6.

If c satisfies Al to A5, but not A6, we shall call it a quasi-regular c-function (not in [Prob j). In this case, c_t is 0 for some Z_i; therefore, even in L_N, $c(h, e)$ cannot always be represented as $c_t(e \cdot h)/c_t(e)$.

(Example: *the straight rule*, [Prob.] p. 227.)

The following theorems are provable on axioms Al to A5; hence they hold for all regular or quasi-regular c-functions.

T2. $c(h, e) \times c(i, e \,.\, h) = c(i, e) \times c(h, e \,.\, i)$. (From A4, A5.).

T3. General division theorem, in two forms.

a. If $c(i, e) > 0$, then $c(h, e \,.\, i) = \dfrac{c(h,e) \times c(i, e.h)}{c(i,e)}$ (From T2.)

b. If $c(i, e) > 0$ and $c(h, e) > 0$, then $\dfrac{c(h, e.i)}{c(h,e)} = \dfrac{c(i, e.h)}{c(i,e)}$ (From (a).)

The fraction on the left-hand side of the equation is known as the *relevance quotient*; the numerator is the *posterior confirmation* of *h* and the denominator is the *prior confirmation* of *h*.

T4. *Special division theorem*. Suppose that *c*(*i*, *e*) > 0, *c*(*h*, *e*) > 0, and *c* (*i*, *e* . *h*) = 1 (*i* is predictable or explainable by *h*). Then

$$\frac{c(h, e \cdot i)}{c(h, e)} = \frac{1}{c(i, e)} \text{ (From T3b).}$$

See the explanations and examples for these theorems in [Prob.] §§ 60 and 61.

VII. *Coherence*

Informal explanation. Let *X* be willing to accept any system of bets in which the betting quotients are equal to the values of a function *c*. If there were a betting system such that *X* would suffer a loss in every logically possible case, *c* would obviously be unsuitable. If there is no such betting system, we shall call *c coherent* (Ramsey, De Finetti). If, moreover, there is no betting system such that *X* would lose in at least on possible case and would not gain in any possible case, we shall call *c strictly coherent* (Shimony).

We assume for the following definitions that L is an interpreted language, that *e* and *h* are sentences of L, that *e* is not L-false, that *c* is a function whose value for any *h*, *e* is a real number, and that *q* and *S* are real numbers (and likewise for e_i, h_i, q_i, S_i).

We represent a bet (of the person *X*) on *h*, given *e*, in language L, with the betting quotient *q* and the total stake *S* as the ordered quintuple < L, *h*, *e*, *q*, *S*> (without reference to *X*):

Dl. *B* is a *bet* $=_{Df}$ for some L, *h*, *e*, *q*, and *S*, *B* = < L, *h*, *e*, *q*, *S*>.

We represent a betting system BS based on the assumption *k* and comprising the bets B_1, B_2, ..., B_n in language L, in accordance with *c* (i.e., the betting quotients are determined by the values of *c*) as the ordered quadruple <{B_1, ..., B_n}, L, *k*, *c*>:

D2. BS is a *betting system* $=_{Df}$ BS = <*K*, L, *k*, *c*>, where *K* = {B_i} (*i* = 1, ..., *n*), B_i = <L, h_i, e_i, q_i, S_i>, *k* is a non-L-false sentence in L, each e_i is either *k* or a conjunction containing *k* as a component and is not L-false, and $q_i = c(h_i, e_i)$.

If X regards a bet on h, given e, with betting quotient q as fair, then he is willing to make a corresponding bet on either side, i.e., either for h or against h. If e is true, the gains are as follows (with $S>0$):

		Gain	
	h	for h	against h
(a)	true	$(1-q)\,S$	$-(1-q)\,S$
(b)	false	$-qS$	qS

Thus X's bet against h can be regarded as a bet for h with negative S. Therefore we admit $S \leq\geq 0$; then D3 covers both bets, for h and against h. $g\,(B, j)$ is the gain which X would obtain from his bet B if j were true.

D3. Let B be a bet $\langle \mathtt{L}, h, e, q, S\rangle$. Let j be a non-L-false sentence in $\mathtt{L}$ which L-implies either e or $\sim e$ and L-implies either h or $\sim h$. $g\,(B, j) =_{\mathrm{Df}}$ the value u such that

either (a) $\vdash j \supset e \,.\, h$, and $u = (1-q)\,S$,
or (b) $\vdash j \supset e \,.\, \sim h$, and $u = -qS$,
or (c) $\vdash j \supset \sim e$, and $u = 0$.

We define $G\,(\mathrm{BS}, j)$ as the total gain from the betting system BS which X would obtain if j were true:

D4. Let BS be $\langle\{B_i\}, \mathtt{L}, k, c\rangle$ $(i = 1,\ldots, n)$. Let j be a non-L-false sentence in $\mathtt{L}$ such that, for every i, j L-implies either e_i or $\sim e_i$, and j L-implies either h_i or $\sim h_i$. Then $G\,(\mathrm{BS}, j) =_{\mathrm{Df}} \sum_{i=1}^{n} g\,(B_i, j)$.

Let BS be $\langle\{B_i\}, \mathtt{L}, k, c\rangle$. Let C_{BS} be the class of the conjunctions j such that (1) j contains as components, for each of the sentences $e_1,\ldots, e_n, h_1,\ldots, h_n$, either the sentence itself or its negation but not both, and no other components, and (2) j is compatible with k. These conjunction represent the possible cases on the basis of the assumption k. We shall say that for a given BS *loss is necessary* if, for every conjunction j in C_{BS}, $G\,(\mathrm{BS}, j) < 0$; that *loss is possible* if, for at least one j in C_{BS}, $G\,(\mathrm{BS}, j) < 0$; and that *positive gain is impossible* if, for every j in C_{BS} $G\,(\mathrm{BS}, j) \leq 0$. We shall say that BS is *vacuous* if, for every j, $G\,(\mathrm{BS}, j) = 0$.

D5. c is a *coherent c-function* for L $=_{\mathrm{Df}}$ there is no betting system in L in accordance with c for which loss is necessary (in other words, for every betting system there is a possible outcome without loss).

D6. c is a *strictly coherent c-function* for L $=_{\mathrm{Df}}$ there is no betting system in L in accordance with c for which loss is possible and positive gain is impossible (in other words, for every non-vacuous betting system there is a possible outcome with positive gain).

T1. If c is strictly coherent, it is also coherent.

T2. (*Ramsey*, *De Finetti*.) Every *coherent c*-function satisfies the axioms Al to A5. In other words, if c violates at least one of the axioms Al to A5, then there is betting system in accordance with c for which loss is necessary.

Example for A4. Suppose that c violates A4 in L. Then there are sentences e, h, and h' in L such that

$$c\,(h, e) \times c\,(h', e \,.\, h) - c\,(h \,.\, h', e) \neq 0.$$

Let $c_1 = c\,(h, e)$, $c_2 = c\,(h', e \,.\, h)$, $c_3 = c\,(h \,.\, h', e)$, and let $c_1\, c_2 - c_3 = D$.
We choose the betting system BS = <{B_i}, L, e, c> (i = 1, 2, 3),

TABLE I
Example for A4

	B_i	$g\,(B_i, j)$ for the four conjunctions in C_{BS}					
i	e_i	h_i	S_i	q_i	$e \,.\, h \,.\, h'$	e . h . ~h′	$e \,.\, {\sim}h \,.\, h'$ $e \,.\, {\sim}h \,.\, {\sim}h'$
1	e	h	$\frac{c_2}{D}$	c_1	$\frac{(1-c_1)c_2}{D}$	$\frac{(1-c_1)c_2}{D}$	$\frac{-c_1c_2}{D}$
2	$e \,.\, h$	h'	$\frac{1}{D}$	c_2	$\frac{1-c_2}{D}$	$\frac{-c_2}{D}$	0
3	e	$h \,.\, h'$	$\frac{-1}{D}$	c_3	$\frac{-(1-c_3)}{D}$	$\frac{c_3}{D}$	$\frac{c_3}{D}$
G(BS,j)=					−1	−1	−1

with *e* as *k*, and with e_i, h_i, and S_i as specified in the table below. (See Table I.) By D2, $q_i = c(h_i, e_i)$. The values of *g* are determined by D3, and those of *G* by D4. We find that, for every *j*, *G* = −1. Thus for the chosen BS, loss is necessary. This betting system is described in Table I.

T3. (*Shimony*.) If *c* violates A6, then there is a betting system in accordance with *c* for which loss is possible and positive gain is impossible. Therefore every *strictly coherent c*-function satisfies the axioms Al to A6.

Proof. Suppose that *c* violates A6 in L_N. Then there are sentences *e*, *h* in L_N such that *c* (*h*, *e*) =1 but *e* does not L-imply *h*, hence *e* **.** ~*h* is not L-false. We take a system of one bet < L_N, *h*, *e*, *c* (*h*, *e*), 1>, and again *e* as *k*. The two possible cases *j* are *e* **.** *h* and *e* **.** ~*h*. The gain is 0 in the first case, and −1 in the second. Thus loss is possible and positive gain is impossible. This applies to any quasi-regular *c*-function, e.g. to the *straight rule* (VI).

T2 gives a *validation for the axioms* Al *to* A5, T3 *for* A6. The following theorem shows that an analogous validation is not possible for any further axioms. (The proof for T4 is given by Kemeny in his paper [34].)

T4. (Kemeny) **a**. Every *c*-function in L which satisfies the axioms Al *to* A5, is *coherent* in L .
b. Every *c*-function in L which satisfies the axioms Al *to* A6, is *strictly coherent* in L .

T5. **a**. A *c*-function is coherent if and only if it is regular or quasi-regular.
b. A *c*-function is strictly coherent if and only if it is regular. (From T2, T3, T4, and VI-T1.)

The classification of *c*-functions defined by T4 and T5 can be presented in the form of the following table:

<table>
<tr><td colspan="2">Axioms satisfied</td><td colspan="2">Type of c-function</td></tr>
<tr><td rowspan="2">A1 to A5</td><td>A6</td><td>regular
strictly coherent</td><td rowspan="2">coherent</td></tr>
<tr><td>not A6</td><td>quasi-regular</td></tr>
</table>

VIII. *Symmetrical c-Functions* (A7)

The system Al to A6 is very weak. It determines no value of $c(h, e)$ except 0 or 1 in special cases. For any pair of factual sentences e, h such that e L-implies neither h nor $\sim h$, the system does not exclude any number between 0 and 1 as a value of $c(h, e)$ ([Prob.] T59-5f, see remark on p. 323). Thus additional axioms are needed. A7 is the first of several *axioms of invariance* of $c(h, e)$ with respect to certain transformations of e and h. These axioms represent the valid core of the classical *principle of indifference*. *Axiom of symmetry* (with respect to individuals):

A7. $c\ (h, e)$ is invariant with respect to any permutation of the individuals.

Dl. m-functions and c-functions which satisfy A7 are said to be *symmetrical* (with respect to individuals). (See [Prob.] §§ 90, 91.)

Read the definitions and explanations of the following concepts in [Prob.] : Ch. III: *division* (D25-4), *isomorphic* sentences (D26-3) and isomorphic *state-descriptions* (§ 27), *individual* and *statistical distributions* (D26-6), *structures* (§ 27) and *structure-descriptions* (*Str*, D27-1), *Q-predicates* (§ 31) and *Q-numbers* (§ 34).

Henceforth it is assumed, unless the contrary is stated, that c satisfies Al to A7 and hence is regular and symmetrical. m is c_t; hence c is based on m.

T1. Let e be isomorphic to e', and h to h'.
a. $c\ (h, e) = c\ (h', e')$. (From A7).
b. $m\ (h) = m\ (h')$. (From (a)).

T2. Let i be an individual distribution for n given individuals with respect to the division $M_1,..., M_k$, with the cardinal numbers $n_1,..., n_k$.
a. The numbers of the individual distributions for the same n individuals which are isomorphic to i is

$$\zeta_i = \frac{n!}{n_1!...n_k!}$$ ([Prob.] T40-32b.)

b. Let j be the statistical distribution corresponding to i. Then $m(j) = \zeta_i \times m(i)$. (From T1b).

T3 is a special case of T2.

T3. Let L_N be a language with N individual constants and k Q-predicates. Let Z_i be a state-description in L_N with the Q-numbers N_j ($j = 1, ..., k$).

a. The number of those state-descriptions in L_N which are isomorphic to Z_i is

$$\zeta_i = \frac{N!}{N_1!\,N_2!\ldots N_k!}\text{. (From T2a.)}$$

b. Let Str_i be the structure-description corresponding to Z_i. Then $m\,(Str_i) = \zeta_i \times m(Z_i)$. (From T2b.)

Therefore a regular and symmetrical m-function for L_N is uniquely determined if we choose as its values for the structure-descriptions in L_N arbitrary positive numbers whose sum is 1. Then, for any Z_i, $m\,(Z_i)$ is determined by T3b and hence the other values by VI-D1c and d.

The subsequent *theorems* T4 *to* T6 *on the direct inductive inference* refer to the following situation. e is a statistical distribution for n given individuals (the 'population') in L_N with respect to the division M_1, M_2 (which is non-M_1) with the cardinal numbers n_1, n_2. $r_i = n_i/n$ (i =1, 2). h is an individual distribution for s of the n individuals (the 'sample') with the cardinal numbers s_1, s_2 ($s_i \leq n_i$). h_{st} is the statistical distribution corresponding to h.

T4. **a**. $c\,(h, e) = \dfrac{\left[{n_1 \atop s_1}\right]\left[{n_2 \atop s_2}\right]}{\left[{n \atop s}\right]}$.

(For $\left[{n \atop s}\right]$, see [Prob.] D40-3.)

b. $c\,(h_{st}, e) = \dfrac{\binom{n_1}{s_1}\binom{n_2}{s_2}}{\binom{n}{s}}$

(For $\binom{n}{m}$, see D40-2.)

c. For given e and s, $c(h_{st}, e)$ has its maximum if s_1/s is equal, or as near as possible, to r_1.
d. For fixed s, let $h_p(p = 0,..., s)$ be the statistical distribution h_{st} with $s_1 = p$ and $s_2 = s - p$. Then

$$\sum_{p=0}^{s} [p \times c\,(h_p, e)] = sr_1\ .$$

e. Let j be a full sentence of 'N_1' with one of the n individual constants in e. Then

$c(j, e) = r_1$.   (For proofs see [Prob.] T94-1.)

We see from T4d that, for given s, the estimate of s_1 on e is sr_1. Hence the estimate of s_1/s is r_1. T4e shows that c for a singular prediction with 'N_1' is r_1. Thus for the direct inference something analogous to the straight rule holds for all symmetrical regular (or quasi-regular) c-functions.

T5. The following holds *approximately* for sufficiently large n, n_1, and n_2. It holds exactly for lim c ($n \to \infty$) if $\lim(n_i/n) = r_i$.

a. $c\,(h, e) = r_1^{s1} \times r_2^{s2}$.

b. Binomial law. $c\,(h_{st}, e) = \binom{s}{s_1} r_1^{s1} r_2^{s2}$.

For proofs and explanations, see [Prob.] § 95.

We shall use the following notations in T6: $\sigma = \sqrt{sr_1r_2}$ ('standard deviation'); $\delta = s_1 - sr_1$ (deviation of s_1 from its estimate); $\phi(u) = (1/\sqrt{2\pi})e^{-u^2/2}$ (the normal function; [Prob.] D40-4a); h_p as in T4d; h' is the disjunction of sentences h_p with p running from $sr_1 - \delta'$ (or the integer nearest to it) to $sr_1 + \delta'$ ($= s_1'$) ; thus h' says that s_1 deviates from its estimate sr_1 to either side by not more than δ', in other words, that s_1/s (the relative frequency of M_1 in the sample) does not deviate from r_1 by more than δ'/s.

T6. The following holds *approximately* for sufficiently large s and n/s.

a. *The normal law*.

$$c(h_{st}, e) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\delta^2/2\sigma^2} = \frac{1}{\sigma}\phi\left(\frac{\delta}{\sigma}\right).$$

b. *Bernoulli-Laplace theorem*

$$c(h', e) = \int_{-\delta'/\sigma}^{+\delta'/\sigma} \phi(u)\mathrm{d}u.$$

c. *Bernoulli's limit theorem*. For fixed r_1 and fixed $q = \delta'/s$, $\lim_{s\to\infty} c(h', e) = 1$.

T6c says the following. If the sample size s increases but a fixed interval $r_1 \pm q$ around the given r_1 is chosen, then $c(h', e)$ (i.e., the probability that the relative frequency of M_1 in the sample lies within the chosen interval) can be brought as near to 1 as desired by making the sample sufficiently large. For explanations and numerical examples, see [Prob.] § 96.

IX. *Estimation*

Read: [Prob.] § 98 about the present situation of the problem of estimation.

Definition of the general estimate function.

Suppose that, on the basis of e, the magnitude u has n possible values: $u_1,..., u_n$. Let h_i say that u has the value u_i (i = 1, ..., n). The *c-mean estimate* of u is the weighted mean of the possible values, with their c-values as weights:

DI. $est(u, e) = \mathrm{Df} \sum_{i=1}^{n} [u_i \times c(h_i, e)].$

T1. A and B are arbitrary fixed constants.
a. $est\ (Au, e) = A \times est\ (u, e)$.
b. $est\ (u + B, e) = est\ (u, e) + B$.
c. $est\ (Au + B, e) = A \times est\ (u, e) + B$.
([Prob.] T100-3,4, and 5).

Analogous results do not generally hold for a non-linear function of u. For example, in general $est(u^2, e) \neq est^2(u, e)$. This leads to a *paradox* in the practical application of estimates ([Prob.] § 100 C). The paradox is eliminated if the rule for the determination of a decision refers to the estimate of only one magnitude, e.g., the gain or the utility resulting from an action.

Truth frequency. Let K be a class of s sentences $i_1,..., i_s$. Let $tf(K)$ be the truth-frequency in K, i.e., the number of true sentences in K. Let $rtf(K)$ be the relative truth-frequency in K, i.e., $tf(K)/s$.

T2. **a**. $est\,(tf, K, e) = \sum_{n=1}^{s} c\,(i_n, e)$. (For this proof. see [Prob.] T104-2a.)

b. $est\,(rtf, K, e) = \frac{1}{s}\sum_{n=1}^{s} c\,(i_n, e)$. (From (a). T1a.)

c. If all sentences in K have the same c-value on e, then the estimate of $rtf(K)$ is equal to this c-value. (From (b).)

The frequency of a property of individuals. Let K be a class of n individuals defined by enumeration. Let $af(M, K)$ be the absolute frequency of M in K, and $rf(M, K)$ the relative frequency, i.e., of $(M, K)/n$. Let K' be the class of the full sentences of M with those individual constants which designate the individuals in K Then

$$af\,(M, K) = tf\,(K') \qquad \text{and} \qquad rf\,(M, K) = rtf\,(K').$$

Therefore the results T2 on estimates of tf and rtf can now be applied to estimates of af and rf.

Direct estimation of frequency. This is based on the direct inference (see VIII-T4). Let e, n, M_1, n_1, r_1, s, and s_1 be as before (VIII-T4). Thus e says that the rf of M_1 in the population is r_1. Let K be the class of the s individuals of the sample.

T3. **a**. $est\,(af, M, K, e) = sr_1$. (From VIII-T4d.)
b. $est\,(rf, M, K, e) = r_1$. (From (a), T1a.)

Predictive estimation of frequency. Here the estimate depends on the chosen c-function. Let e be any non-L-false sentence, h a full sentence of M for a new individual, and K any finite, non-empty class of new individuals.

T4. $est\ (rf, M, K, e) = c\ (h, e)$. (From T2c.)

Thus the confirmation of a singular prediction with M is equal to the estimate of rf of M. This relation was used earlier for an informal explanation of inductive probability ([Prob.] § 41D).

X. *The Functions* $e^{\dagger}$ *and* c^{*}

In discussions on the principle of indifference, some authors have proposed to give equal *a priori* probabilities to all individual distributions (for a given domain of individuals and a given division of properties). Other authors have proposed the same for all statistical distributions. In our terminology, the controversy concerns the choice of one of the following two rules:

(A) All individual distributions have equal m-values.
(B) All statistical distributions have equal m-values.

However, each of these rules leads to contradictions if applied to different divisions (see the examples in [Continuum] p. 39).

Each of the rules becomes consistent if it is restricted to one division (for a given finite language), viz. the division of the Q-predicates, as follows:

(A′) All state-descriptions have equal m-values.
(B′) All structure-descriptions have equal m-values.

The function $c^{\dagger}$. There is exactly one symmetrical, regular m-function which fulfills (A′), viz. $m^{\dagger}$ defined by Dl.

Let L_N be a language with N individual constants and k Q-predicates.

T1. **a**. The number of state-descriptions in L_N is $\zeta_N = k^N$. ([Prob.] T40-31c.)
b. The number of structure-descriptions in L_N is

$$\tau_N = \binom{N+k-1}{k-1} = \frac{(N+k-1)!}{N!(k-1)!}.$$ ([Prob.] T40-33b.)

Let Z_N be any state-description in L_N with the Q-numbers $N_1, \ldots, N_k$. We define:

D1. $$m^{\dagger}(Z_N) =_{\mathrm{Df}} \frac{1}{k^N}$$

D2. $c^{\dagger}(h,e) =_{\text{Df}} \dfrac{m^{\dagger}(e.h)}{m^{\dagger}(e)}$.

Let e_N be an individual distribution for any N individuals for the division of the k Q-predicates with the same Q-numbers $N_1, \ldots, N_k$ (the same as in Z_N). Let h_j be a full sentence of Q_j for a new individual.

T2. **a**. $m^{\dagger}$ is regular and symmetrical. (From D1.)

b. $m^{\dagger}(e_N) = \dfrac{1}{k^N}$. (From D1, since e_N is isomorphic to Z_N.)

c. $c^{\dagger}(h_j, e_N) = 1/k$.

Proof. $e_N . h_j$ is isomorphic to a state-description in L_{N+1}, hence $m^{\dagger} = 1/k^{N+1}$ (from Dl). The result is obtained by D2 and (b).

T2c shows that *et (h;, e*$_N$*)* is independent of e_N. It violates the principle of learning from experience and hence is unacceptable ([Prob.] p. 565). However, this function was proposed by C. S. Peirce, Keynes, and Wittgenstein.

The function c*. There is exactly one symmetrical, regular *m*-function which fulfills (B′), viz. *m** defined by D3.

D3. $m^*(Z_N) =_{\text{Df}} \dfrac{1}{\tau_N \zeta_i}$

$$= \frac{N_1!\ldots N_k!(k-1)!}{(N+k-1)}.\text{(From T1b, VIII-T3a.)}$$

T3. **a**. For any structure-description in $\mathrm{L}N$, $m^* = \dfrac{1}{\tau_N}$. Thus *m** fulfills (B). (From VIII-T3b.)

b. *m** is regular and symmetrical. (From D3.)

c. $m^*(e_N) = \dfrac{N_1!\ldots N_k!(k-1)!}{(N+k-1)}$. (From D3.)

*c** is based on *m**:

D4. $c^*(h, e) =_{\text{Df}} \dfrac{m^*(e.h)}{m^*(e)}$.

T4. $c^*(h_j, e_N) = \frac{N_j + 1}{N + k}$.

Proof. $e_N \cdot h_j$ is isomorphic to a state-description in $\mathfrak{L}_{N+1}$ with the Q-numbers $N_1,..., N_j + 1, . . ., N_k$. Therefore its m^*-value is like that of e_N in T3c, but with $N_j + 1$ instead of N_j and $N + 1$ instead of N. Hence the result by D4.

Let M be a disjunction of w Q-predicates ($0 < w < k$) and N_M be the sum of the Q-numbers of these Q-predicates in e_N. Hence w is the *logical width* of M ([Prob.] § 32). Let h_M be a full sentence of M for a new individual.

T5. $c\,(h_m, e_N) = \frac{N_M + w}{N + k}$. (From T4 and A3.)

Consider a sequence of samples of increasing size N but such that $r = N_m/N$ remains constant. Then the value of $c^*\,(h_m, e_N)$ moves from w/k (for $N = 0$, i.e., tautological evidence) towards r (which is the limit for $N \rightarrow \infty$).

For further explanations and theorems on c^* see [Prob.] § 110.

XI. *Further Axioms of Invariance* (A8-A11)

A8. $c\,(h, e)$ is invariant with respect to any *permutation of the predicates* of any family.

T1. Let F be a family of k primitive predicates 'P_1', …, 'P_k'. Let $h_1,..., h_k$ be full sentences of these predicates with the same individual constant, and h be the disjunction of these sentences.

a. (Lemma.) For any e, $c\,(h, e) = 1$. (From A2, since h is L-true.)

b. Suppose that e' does not contain any predicate of F. Then for any i (= 1, ..., k), $c(h_i, e') = 1/k$.

Proof. The k values $c(h_i, e')$ are equal (by A8). Their sum = $c(h_i, e')$ (by A3) = 1 (by (a)). Hence the assertion.

c. $m(h_i) = 1/k$. (From (b).)

A9. $c\,(h, e)$ is invariant with respect to any *permutation of families of the same size*.

A10. For non-general h and e, $c(h, e)$ is *independent of the total number of individuals*. (A10 corresponds to the requirement of a fitting c-sequence, [Prob.] § 57C.)

A11. $c(h, e)$ is *independent of the existence of other families* than those occurring in h or e.

XII. *Learning from experience* (A12)

The intuitive *principle of learning from experience* says that, other things being equal, the more frequently a kind of event has been observed, the more probable is its occurrence in the future. This is expressed more exactly in the *axiom of instantial relevance* (first proposed in Carnap [16])

A12. Suppose that e is non-L-false and non-general, and i and h are full sentences of the same factual, molecular predicate 'M' with distinct individual constants not occurring in e.

a. $c(h, e \,.\, i) < c(h, e)$. ****THE '<' SYMBOL SHOULD HAVE A VERTICAL LINE THRU IT****

b. $c(h, e \,.\, i) \neq c(h, e)$.

Both $c^{\dagger}$ (X) and the straight rule (VI) fulfill part (a) of A12, but violate part (b). With $c^{\dagger}$, i is always irrelevant for h. With the straight rule, i is irrelevant for h if e is a conjunction of full sentences of 'M'; in this case both c-values are 1.

T1. Let e, i, h, and M be as in A12.

a. $c(h, e \,.\, i) > c(h, e)$; i is positively relevant for h on e.

b. Let j be a conjunction of n full sentences of 'M' ($n \geq 2$) with n distinct individual constants which do not occur in e or h. Then

$c(h, e \,.\, j) > c(h, e)$. (From (a).)

c. $c(h, e \,.\, {\sim}i) < c(h, e)$; ~i is negatively relevant for h on e. (From (a) and [Prob.] T65-6e.)

d. $c(h, e, i) > c(h, e \,.\, {\sim}i)$. (From (a), (c).)

XIII. *The language L_F with one family F* (Al 3)

This and the subsequent sections refer to a language L_F whose primitive predicates are k predicates 'P_1',…, 'P_k' of a family F ($k \geq 2$). A sentence in L_F may contain any number of individual constants but no variables. e_F is an individual distribution for s individuals with respect to F with the cardinal numbers s_i ($i = 1,\ldots, k$). $h_l,\ldots, h_k$ are full sentences of 'P_1', …, 'P_k', respectively, with the same individual constant, which does not occur in e_F.

A13. Meaning postulates for F:
a. $\vdash h_1 \vee h_2 \vee \ldots \vee h_k$.
b. If $i \neq j$, $h_i \,.\, h_j$ is L-false.

m (e_F) is independent of other individuals (A10) and other families (A11). It depends not on the particular individuals in e_F but only on their numbers s_i. Therefore:

T1. For any m-function m fulfilling the axioms, there is, for any k, a representative mathematical function M_k of k arguments such that, for any e_F,

$$m\,(e_F) = M_k\,(s_1, s_2, \ldots, s_k).$$

T2. M_k is invariant with respect to any permutation of the k arguments. (From A8.)

$e_F \,.\, h_1$ is an individual distribution for $s + 1$ individuals with the cardinal numbers $s_1 + 1$, $s_2, \ldots, s_k$. We define:

D1. $C_k\,(s_1; s_2, \ldots, s_k) =_{\mathrm{Df}} \dfrac{M_k\,(s_1+1, s_2, \ldots, s_k)}{M_k\,(s_1, s_2, \ldots, s_k)}$.

T3. **a**. For any c-function c and any k, there is a *representative mathematical function* C_k of k arguments such that, for any e_F, $c\,(h_i, e_F) = C_k\,(s_1; s_2, \ldots, s_k)$. Analogously for h_2, etc.
b. C_k is invariant with respect to any permutation of the k-1 arguments following the first.

I shall sometimes write 'M' and 'C' without subscripts.

T4. For any k numbers $n, p, s_3, \ldots, s_k$ whose sum is s, the following holds. ('$---$' stands for '$s_3, \ldots, s_k$' ; this expression drops out if $k = 2$; in this case $n + p = s$.)

$$\frac{C(n; p+1,---)}{C(p; n+1,---)} = \frac{C(n; p,---)}{C9p; n,---)}$$

(Here $p + n + 1 + ––– = s + 1$.)

Proof. The following holds identically:

$$\frac{M(n+1, p+1, ---)}{M(n, p+1, ---)} \times \frac{M(n, p+1, ---)}{M(n, p, ---)} = \frac{M(n+1, p+1, ---)}{M(n+1, p, ---)} \times \frac{M(n+1, p, ---)}{M(n, p, ---)}$$

According to D1, the first quotient is $C(n; p + 1, –––)$; the second is (by T2) equal to

$$\frac{M(p+1, n, ---)}{M(p, n, ---)} = C(p; n, ---);$$

the third becomes (again with reordering of arguments) $C(p; n + 1, –––)$, and the fourth $C(n; p, –––)$. Hence the theorem.

T5. **a**. $\sum_{i=1}^{k} c\,(h_i, e_F) = 1$. (From A13a.)

b. $\sum_{i=1}^{k} C\,(s_i; s_1, ..., s_{i-1}, s_{i+1}, ..., s_k) = 1$. (From (a).)

XIV. *The Axiom of Predictive Irrelevance* (A14)

Let e_1 be formed from e_F by replacing each predicate except 'P_1' with '$\sim P_1$'. Hence e_1 is an individual distribution for the s individuals with respect to the division P_1, $\sim P_1$, with the cardinal numbers s_1 and $s - s_1$. e_2, ..., e_k are formed analogously.

For given k, $c(h_1, e_1)$ depends only on s_1 and s. It can therefore be represented by a function $G_k(s_1; s)$. Analogously for $i = 2, ..., k$ (by A8).

T1. For any c-function c and any k, there is a *representative mathematical function* G_k such that, for $i = 1, ..., k$.

$c\,(h_i, e_i) = G_k\,(s_i; s)$.

T2. Suppose that $s_1 < s$. Let e'_1 be like e_1 but with the cardinal numbers $s_1 + 1$ and $s - s_1 - 1$.

a. $c(h_1, e'_1) > c(h_1, e_1)$. (From XII-Tld.)

b. $G_k(s_1 + 1; s) > G_k(s_1; s)$ (From (a).)

The axiom of predictive irrelevance says that of the k cardinal numbers in e_F all except s_1 are irrelevant for h_1

A14. For $k > 2$, $c(h_1, e_F) = c(h_1, e_1)$.

This axiom is not a necessary condition for the adequacy of c. But it is a customary (usually tacit) assumption, and it leads to a great simplification of the system. If $k = 2$, then e_1 is the same as e_F and therefore A14 is fulfilled trivially.

T3. For any k (≥ 2) any any i:
a. $c(h_i, e_F) = c(h_i, e_i)$. (From A14, A8.)
b. For any numbers $s_2, ..., s_k$ whose sum is $s - s_i$, $C_k(s_1; s_2, ..., s_k) = G_k(s_1; s)$. (From (a).)

I shall often write 'G' for 'G_k'.

T4. For any sequence of k numbers $s_1,..., s_k$ whose sum is s,
$$\sum_{i=1}^{k} G(s_i; s) = 1. \text{ (From XIII-T5.)}$$

Special cases of T4:

T5. **a**. $G(s; s) + (k - 1) G(0; s) = 1$. (From T4 for the sequence s, 0, ..., 0.)
b. $G(s + 1; s + 1) = 1 - (k - 1) G(0; s + 1)$. (From (a).)
c. $G(1; 1) = 1 - (k - 1) G(0; 1)$. (From (a).)
d. $G(s; s + 1) + G(1; s + 1) + (k - 2) G(0; s + 1) = 1$. (Sequence s, 1, 0,..., 0.)

The following development has the aim to show (1) that, if all values of G for s are given, the values for $s + 1$ are uniquely determined, and (2) if $G(0; 1)$ is given, all values of G are uniquely determined. For these results it is *presupposed that* $k > 2$.

T6. For $k > 2$; for any n, p, s such that $n + p \leq s$.
a. $\dfrac{G(n; s+1)}{G(p; s+1)} = \dfrac{G(n; s)}{G(p; s)}$. (From XIII-T4.)
b. $G(n; s + 1) = G(0; s + 1) \dfrac{G(n; s)}{G(0; s)}$ (From (a) with $p = 0$.).

T7. For $k > 2$.

$$G(0; s+1)\left[\frac{G(s;s)}{G(0;s)}+\frac{G(1;s)}{G(0;s)}+k-2\right]=1.$$

(From T5d, by transforming the first two of its *G*-terms according to T 6b.)

Now aim (1) has been reached. If all *G*-values for *s* are given, $G(0; s + 1)$ is determined by T7, then the values $G(n; s + 1)$ for $n = 1, \ldots, s$ are determined by T6b, and $G(s + 1; s + 1)$ by T5b. Thus all values for $s + 1$ are determined.

We have also attained aim (2). If $G(0; 1)$ is given, $G(1; 1)$ is determined by T5c. These are all the *G*-values for $s = 1$. They determine the values for $s = 2$, and so on. Thus all *G*-values are determined by $G(0; 1)$. The following theorem gives the explicit form.

T8. For $k > 2$, for any *s* and *n* ($0 \leq n \leq s$),

$$G(n; s) = \frac{n-(kn-1)G(0;1)}{s-(s-1)kG(0;1)}.$$

(This can be proved by mathematical induction with respect *to* s. (1) The theorem holds for $s = 1$ (for $n = 0$ it holds identically, for $n = 1$ by T5c). (2) If the theorem holds for a given *s*, it holds likewise for $s + 1$; this can be shown with the help of the theorems T7, T6b, and T5b, which determine the *G*-values for $s + 1$ on the basis of those for *s*. Hence the theorem holds for every *s*.)

Suppose that the value of $G(0; 1)$ has been chosen. Then all values of *G* can be determined. The following theorem T9 shows that the *m*-value of any state-description is determined by the value of *G*. Thereby the *m*-values for all sentences and the *c*-values for all pairs of sentences are determined (see VI).

T9. Let Z_F be a state-description for *N* individuals and for the *k* predicates of the family *F*, with the cardinal numbers N_i ($i = 1, \ldots, k$). Then

$$m(Z_F) = \prod_i \prod_{n=0}^{N_i-1} G(n; S_i + n),$$

where $\prod_i$ runs through those values of i for which $N_i > 0$;

$S_i = \sum_{h=1}^{i-1} N_h$, $S_1 = 0$. (For the proof see [Continuum] § 5.)

XV. *The λ-system* (Al 5)

We shall construct a system of all *c*-functions fulfilling our axioms. We call it the λ -system, because the *c*-functions will be characterized by the values of a parameter λ.

We have seen that, for a given $k(> 2)$, all values of G are determined by G (0; 1). The latter value can be freely chosen within certain boundaries. We shall now determine these boundaries.

From XI-T1c:

(1) $c(h_i, t) = 1/k$.

Hence with XII-T1c (based on A12):

(2) $G(0; 1) < 1/k$.

If we were to choose $G(0; 1) > 1/k$, then c would violate not only A12b but also A12a and therefore be unacceptable. If we choose $G(0;1) = 1/k$, then only A12b is violated. This c does not belong to the λ -system, but will nevertheless be discussed as a boundary case (we shall find that it is the same as $c^\dagger$ in X).

The following is obvious (from Al):

(3) $G(0; 1) \geq 0$.

If we choose $G(0; 1) = 0$, then the resulting c fulfills Al to A5, but violates A6. Hence it is quasi-regular. Therefore it does not belong to the λ -system. It will, however, be discussed as a boundary case; we shall see that it corresponds to the straight rule.

T1. $0 < G(0; 1) < 1/k$.

It can also be shown that, if any value between 0 and $1/k$ is chosen for $G(0; 1)$, then the resulting *c*-function fulfills our axioms.

We define first an auxiliary parameter:

D1. $\lambda_k''' =_{\text{Df}} kG_k\,(0; 1)$.

From T1:

T2. $0 < \lambda_k''' < 1$.

I shall usually write 'λ'''' for 'λ_k''''. From XIV-T8:

T3. For $k > 2$, for any s and n ($0 \le n \le s$).

$$G(n; s) = \frac{n-(n-1/k)\lambda'''}{s-(s-1)\lambda'''}$$

We shall mostly use, not λ''', but $\lambda = \lambda'''/(1-\lambda''')$. The use of λ leads to a simpler formula for $G(n; s)$ (T4c). However, in the case of $G(0; 1) = 1/k$, $\lambda''' = 1$, while λ is infinite. Therefore in this case λ is less convenient than λ'''. But this case is not included in our system.

D2. $\lambda_k =_{\text{Df}} \dfrac{kG_k(0;1)}{1-kG_k(0;1)}$.

T4. **a**. $\lambda''' = \lambda/(\lambda+1)$.

b. $G(0; 1) = \dfrac{\lambda}{k(\lambda+1)}$.

c. For $k > 2$, $G(n; s) = \dfrac{n+\lambda/k}{s+\lambda}$. (From T3, (a).)

The case $k = 2$. The important results at the end of XIV can be proved only if $k > 2$. (This seems surprising, since A14, on which the results are based, holds also for $k = 2$.) *A new axiom must be added for* $k = 2$. T4c shows that, for given $k(> 2)$, s, and $G(0; 1)$, $G(n; s)$ is a linear function of n. We assume as an axiom that the same holds for $k = 2$:

A15. For given s and $G_2(0; 1)$, $G(n; s)$ is a linear function of n.

Note that $G_2(n; s) = C_2(n; s-n)$ (see XVI-T3b). Therefore we have (without use of A15):

(4) $G_2(n; s) + G(s-n; s) = 1$. (From XIII-T5b.)

(5) $$\frac{G_2(n;s+1)}{G_2(s-n;s+1)} = \frac{G_2(n;s)}{G_2(s-n;s)}.$$ (From XIII-T4.)

From (4) and (5) we can derive with A15:

T5. $G_2(n; s+1) = G_2(0; s+1)\dfrac{G_2(n;s)}{G_2(0;s)}$.

This corresponds to XIV-T6b. Then, in analogy to the earlier proofs, we can now prove the analogues of T3 and of T4c for $k = 2$; the latter is

T6. $G_2\,(n;\, s) = \dfrac{n+\lambda/2}{s+\lambda}$.

For any λ, let c_λ be the c-function characterized by λ, and m_λ be the corresponding m-function. Using our result for $G(n;\, s)$ (T4c, T6), we obtain T7a from XIV-T9 (for the proofs of T7b and c and the notation $\left[{r \atop n}\right]$, see [Continuum] § 10).

T7. **a**. $m_\lambda(Z_F) = \prod\limits_{\substack{i \\ (N_i>0)}} \prod\limits_{n=0}^{N_i-1} \dfrac{n+\lambda/k}{S_i+n+\lambda} =$

b. $$= \frac{\prod\limits_i \left[\left(\frac{\lambda}{k}\right)\left(\frac{\lambda}{k}+1\right)\left(\frac{\lambda}{k}+2\right)\ldots\left(\frac{\lambda}{k}+N_i-1\right)\right]}{\lambda(\lambda+1)(\lambda+2)\ldots(\lambda+N-1)}$$

c. $$= \frac{\prod\limits_i \begin{bmatrix} N_i+\lambda/k-1 \\ N_i \end{bmatrix}}{\begin{bmatrix} N_i+\lambda-1 \\ N \end{bmatrix}}.$$

d. Let Str_F be the structure-description corresponding to Z_F. Then

$$m_\lambda(Str_F) = \frac{\prod\limits_i \begin{bmatrix} N_i+\lambda/k-1 \\ N_i \end{bmatrix}}{\begin{bmatrix} N_i+\lambda-1 \\ N \end{bmatrix}}.$$

(From (c) with VIII-T3.)

XVI. *Various c-Functions in the λ-System*

Let the predicate 'M' be defined as a disjunction of w predicates of F: $P_1 \vee P_2 \vee \ldots \vee P_w$. Hence its logical width is w. Let e_F be as before, and h_M a full sentence of 'M' for a new individual. Let e_M be $s_1 + s_2 + \ldots + s_w$. Then we have (from XV):

(1) $c_\lambda\,(h_M,\, e_F) = \dfrac{s_M + w\lambda/k}{s+\lambda}$.

This is

$$\frac{\frac{s_M}{s}s+\frac{w}{k}\lambda}{s+\lambda},$$

thus it is the weighted mean of the observed relative frequency s_M/s and the relative width w/k, with weights s and λ, respectively. s_M/s is an empirical factor in the situation, and w/k is a logical factor. λ is thus the weight of the logical factor. The greater the chosen λ, the closer to w/k is the above c-value.

Example. For an even k, we take a predicate 'M' with $w = k/2$. In e_F, let $s = 10$, $s_M = 1$. Then the c_λ-value in (1), for various choices of λ, is as follows (see [Continuum] (12-19)):

λ =	0	1	2	4	8	16	32	∞
$c_\lambda(h_M,e_F)$=	0.1	0.136	0.167	0.214	0.278	0.346	0.405	0.5

For $\lambda = 0$, $c = s_M/s$. This is the straight rule, which violates A6. ([Continuum] § 14.)

For $\lambda = \infty$, $c = w/k = c\,(h_M, t)$. This is $c^\dagger$, which violates A 12b. ([Continuum] § 13.)

These are the two extreme methods, not included in the λ -system. In this system, we take $0 < \lambda < \infty$; hence the above c is between s_M/s and w/k (if these are unequal).

For families of different sizes (each in a separate language) we distinguish two kinds of inductive methods.

Inductive methods of the first kind: a fixed value is chosen for λ, independent of k. ([Continuum] § 11.)

Inductive methods of the second kind: λ_k is dependent upon k. The simplest form is: $\lambda_k = C_k$, with a constant C. The simplest method of this form takes $C = 1$, hence $\lambda_k = k$; thus from (1):

$$(2) \qquad c_\lambda(h_M, e_F) = \frac{s_M + w}{s+k}.$$

This is the function c^*(see X).

XVII. *A Language with Two Families* (A16)

The language L contains two families: F^1 consists of k_1 predicates: 'P_1^1' 'P_2^1', etc; and F^2 of k_2 predicates: P_1^2 'P_2^2' etc. There are

$k = k_1k_2$ Q-predicates; Q_{ij} is the conjunction $P_i^1 \,.\, P_j^2$ ($i = 1, ..., k_1$: $j = 1,\ldots, k_2$).

Let e^1 be an individual distribution for F^1, and e^2 for F^2, both for the same s individuals. Let e be $e^1 \,.\, e^2$. This is an individual distribution for the k Q-predicates; let s_{ij} be the number of individuals with Qij.

We take the same λ for both families. Then we can determine $m_\lambda(e^1)$ and $m_\lambda(e^2)$ (by XV-T7).

Problem: What is to be taken as value of $m_\lambda(e)$? This is not determined by the previous axioms. We shall now consider two attempts at a solution, and then take a combination of them.

First tentative solution. We take the class of the k Q-predicates as the *pseudo-family* $F^{1,2}$. Then we define $m^{1,2}$ for $F^{1,2}$, as if the latter were a real family; hence, in analogy to XV-T7c

D1. $$m_\lambda^{1,2}(\mathrm{e}) =_{\mathrm{Df}} \frac{\prod_{i=1}^{k_1}\prod_{j=1}^{k_2}\left[\begin{matrix} s_{ij}+\frac{\lambda}{k}-1 \\ s_{ij}\end{matrix}\right]}{\left[\begin{matrix} s+\lambda-1 \\ s\end{matrix}\right]}.$$

$m_\lambda^{1,2}(\mathrm{e})$ depends only on the Q-numbers s_{ij} in e, not on the P-numbers in e^1 or e^2.

Second tentative solution. We define $m^{1\|2}(e)$ as the product of the m-values for the two families separately:

D2. $$m_\lambda^{1\|2}(e) =_{\mathrm{Df}} = m_\lambda(e^1) \times m_\lambda(e^2) .$$

$m^{1\|2}(e)$ depends only on the P- numbers, not on the Q-numbers.

We shall examine the two solutions with the help of the following three examples A, B, C, of individual distributions for $s = 20$ individuals, with $k_1 = k_2 = 2$ (the numerals in the four cells indicate the Q-numbers; the marginal numerals indicate the P-numbers for the two families).

A			F^2	
			P_1^2 10	P_2^2 10
F^1	P_1^1	10	5	5
	P_2^1	10	5	5

B	F^2	
	P_1^2 10	P_2^2 10
10	10	0
10	0	10

C	F^2	
	P_1^2 10	P_2^2 10
20	10	10
0	0	0

The following two *requirements* (or desiderata) I and II seem plausible.

(I) We should have: $m(A) < m(B)$, because B is more uniform than A.

This requirement is satisfied by $m^{1,2}$ (because the Q-numbers are equal in A, unequal in B), but not by $m^{1|2}$ (this has equal values for A and for B, because the P-numbers are the same).

(II) We should have: $m(B) < m(C)$ because the distribution for F^1 is more uniform in C than in B, while that for F^2 is the same in C as in B.

This requirement is in accord with the customary analogy inference (‘horse-donkey inference’). However, it is not satisfied by $m^{1,2}$ (this has equal values for B and for C, because the Q-numbers are the same). It is satisfied by $m^{1|2}$.

Thus both solutions are unsatisfactory. Generally, any solution that uses only the P-numbers cannot satisfy I, and any solution that uses only the Q-numbers cannot satisfy II. An adequate solution must use both the P-numbers and the Q-numbers. This is done in the third solution, which satisfies both requirements.

Third solution. We define $m_{\lambda,n}(e)$ as a weighted mean of the first two solutions, with the weights η and $1-\eta$, where η is a new parameter

D3. $m_{\lambda,n}(e) =_{\mathrm{Df}} \eta\, m^{1|2}(e) + (1-\eta)\, m^{1|2}(e)$.

The parameter η may be chosen, independently of λ, such that $0 < \eta < 1$. The greater η is, the stronger is the influence by analogy (i.e., the greater is the difference between the two c-values in A16 below). The method can easily be extended to more than two families; no new parameter is needed. (The method was worked out in collaboration with John Kemeny.)

The requirement II can be represented in a generalized form as follows:

A16. *Axiom of analogy*. Let e be an individual distribution for two families (with any k_1 and k_2). Let i and j be full sentences of Q_{11} and Q_{12}, respectively, with the same individual constant not occurring in e. Let h be a full sentence of Q_{12} with another individual constant not occurring in e. Then

$$c(h, e \,.\, i) > c(h, e \,.\, j).$$

The generalization for other Q-predicates follows by A8.

XVIII. *An Infinite Domain of Individuals* (A17)

Let the domain of L_N contain N individuals, and that of L_∞ be denumerably infinite. According to A10, the values of c for non-general sentences are in L_∞ the same as in L_N. If either e or h or both contain variables, a new axiom is needed. We take the value of c in L_∞ as the limit of its values in finite languages (see [Prob.] § 56)

A17. *Axiom of the infinite domain*. Let ${}_N c$ be a c-function for L_N . Then the corresponding c-function ${}_\infty c$ for L_∞ is determined as follows:
$${}_\infty c\ (h, e) = \lim_{N\to\infty} {}_N c(h, e).$$

BIBLIOGRAPHY

A list of selected publications, almost all since 1950. For earlier publications, see the bibliography in Carnap [Prob.].

[1] Bar-Hillel, Y., ‘A Note on State-Descriptions’, *Philosophical Studies* **2** (1951) 72-75.

[2] Bar-Hillel, Y., ‘A Note on Comparative Inductive Logic’, *British Journal for the Philosophy of Science* **3** (1953) 308-310.

[3] Bar-Hillel, Y., ‘Comments on ‘Degree of Confirmation’ by Prof. K. R. Popper’, *British Journal for the Philosophy of Science* **6** (1955) 155-157.

[4] Bar-Hillel, Y., ‘Further Comments on Probability and Confirmation: A Rejoinder to Prof. Popper’, *British Journal for the Philosophy of Science* **7** (1956) 245-248.

[5] Bar-Hillel, Y. and Carnap, R., ‘Semantic Information’, *British Journal for the Philosophy of Science* **4** (1953), 147-157. Repr. in *Language and Information* (by Y. Bar-Hillel).

[6] Black, Max, ‘The Justification of Induction’, *Language and Philosophy*, Cornell University Press, Ithaca, 1949.

[7] Braithwaite, R. B., *Scientific Explanation*, Cambridge University Press, Cambridge, 1953.

[8] Burks, Arthur W., ‘Reichenbach’s Theory of Probability and Induction’, *Review of Metaphysics* **4** (1951) 377-393.

[9] Burks, Arthur W., ‘The Presupposition Theory of Induction’, *Philosophy of Science* **20** (1953) 177-197.

[10] Burks, Arthur W., ‘On the Significance of Carnap’s System of Inductive Logic for the Philosophy of Induction’, *The Philosophy of Rudolf Carnap* (ed. by P. A. Schilpp), Open Court Publ. Co., La Salle, 1963, pp. 739-760.

[11] Carnap, Rudolf, [Prob.] *Logical Foundations of Probability*, The University of Chicago Press, Chicago, 1950.

[12] Carnap, Rudolf, *The Nature and Application of Inductive Logic*, consisting of six sections from *Logical Foundations of Probability*, The University of Chicago Press, Chicago, 1951.

[13] Carnap, Rudolf, 'The Problem of Relations in Inductive Logic', *Philosophical Studies* **2** (1951) 75-80.

[14] Carnap, Rudolf, [Continuum] *The Continuum of Inductive Methods*, The University of Chicago Press, Chicago, 1952.

[15] Carnap, Rudolf, 'Meaning Postulates', *Philosophical Studies* **3** (1952) 65-73.

[16] Carnap, Rudolf, 'On the Comparative Concept of Confirmation', *British Journal for the Philosophy of Science* **3** (1953) 311-318.

[17] Carnap, Rudolf, 'Inductive Logic and Science', *Proceedings of the American Academy of Arts and Sciences* **80** (1953) 189-197.

[18] Carnap, Rudolf, 'Remarks to Kemeny's Paper', *Philosophy and Phenomenological Research* **13** (1953) 375-376.

[19] Carnap, Rudolf, 'What is Probability?', *Scientific American* **189** (1953, September) 128-138.

[20] Carnap, Rudolf, *I. Statistical and Inductive Probability, II. Inductive Logic and Science* (same as above in *Proc. Amer. Acad. Arts & Sciences*), Galois Institute of Mathematics and Art, Brooklyn, N.Y., 1955.

[21] Carnap, Rudolf, 'Remarks on Popper's Note on Content and Degree of Confirmation', Bri*tish Journal for the Philosophy of Science* **7** (1956) 243-244.

[22] Carnap, R. and Bar-Hillel, Y., *An Outline of a Theory of Semantic Information*, Technical Report Number 247, Research Laboratory of Electronics, M.I.T., 1952. Repr. in *Language and Information* (by Y. Bar-Hillel).

[23] Davidson, D., McKinsey, J.C., and Suppes, P., 'Outlines of a Formal Theory of Value', *Philosophy of Science* **22** (1955) 140-160.

[24] De Finetti, Bruno, 'Recent Suggestions for the Reconciliation of Theories of Probability', *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (1951) 217-225.

[25] De Finetti, Bruno, 'La vrai et la probable', *Dialectica* **3** (1949) 79-92.

[26] Feigl, Herbert, 'Scientific Method without Metaphysical Presuppositions', *Philosophical Studies* **5** (1954) 17-29.

[27] Feys, R. (ed.), *Theorie des probabilités*, Exposés sur ses fondements et ses applications, 1953.

[28] Goodman, Nelson, *Fact, Fiction, and Forecast*, Harvard University Press, Cambridge, 1955.

[29] Hutten, Ernest, 'Probability-Sentences', *Mind* **61** (1952) 39-56.

[30] Kemeny, John, 'Extension of the Methods of Inductive Logic', *Philosophical Studies* **3** (1952) 38-42.

[31] Kemeny, John, 'A Contribution to Inductive Logic', *Philosophy and Phenomenological Research* **13** (1953) 371-374.

[32] Kemeny, John, 'The Use of Simplicity in Induction', *Philosophical Review* **62** (1953),391-408.

[33] Kemeny, John, 'A Logical Measure Function', *Journal of Symbolic Logic* **18** (1953) 289-308.

[34] Kemeny, John, 'Fair Bets and Inductive Probabilities', *Journal of Symbolic Logic* **20** (1955) 263-273.

[35] Kemeny, J. and Oppenheim, P., 'Degree of Factual Support', *Philosophy of Science* **19** (1952) 307-324.

[36] Lehman, R. Sherman, 'On Confirmation and Rational Betting', *Journal of Symbolic Logic* **20** (1955) 251-262.

[37] Nagel, Ernest, *Principles of the Theory of Probability*. *International Encyclopedia*

of Unified Science, 1, No. 6, The University of Chicago Press, Chicago, 1939.
[38] Neyman, Jerzy, 'The Problem of Inductive Inference', *Communications On Pure and Applied Mathematics* **8** (1955) 13-45.
[39] Pap, Arthur, *Analytische Erkenntnistheorie*. Kritische Übersicht über die neueste Entwicklung in USA und England. Springer Verlag, Wien, 1955.
[40] Popper, Karl, 'Degree of Confirmation', *British Journal for the Philosophy of Science* **5** (1954) 143-149.
[41] Popper, Karl, 'Two Autonomous Axiom Systems for the Calculus of Probabilities', Br*itish Journal for the Philosophy of Science* **6** (1955) 51-57.
[42] Popper, Karl, 'Content' and 'Degree of Confirmation': A Reply to Dr. Bar-Hillel', Bri*tish Journal for the Philosophy of Science* **6** (1955) 157-163.
[43] Popper, Karl, 'The Demarcation Between Science and Metaphysics', *The Philosophy of Rudolf Carnap* (ed. by P. A. Schilpp), Open Court Publ. Co., La Salle, 1963, pp. 183-226.
[44] Rubin, H. and Suppes, P., 'A Note on Two-Place Predicates and Fitting Sequences of Measure Functions', *Journal of Symbolic Logic* **20** (1955) 121-122.
[45] Savage, Leonard J., *The Foundations of Statistics*, John Wiley, New York, 1954.
[46] Schilpp, Paul A. (ed.), *The Philosophy of Rudolf Carnap* (The Library of Living Philosophers), Open Court Pub]. Co., La Salle, 1963.
[47] Shimony, Abner, 'Coherence and the Axioms of Confirmation', *Journal of Symbolic Logic* **20** (1955) 1-28.
[48] Simon, Herbert A., 'Prediction and Hindsight as Confirmatory Evidence', *Philosophy of Science* **22** (1955) 227-230.
[49] Stegmüller, Wolfgang, 'Bemerkungen zum Wahrscheinlichkeitsproblem', *Studium Generale* **6** (1953) 563-593.
[50] Vietoris, L., 'Zur Axiomatik der Wahrscheinlichkeitsrechnung', *Dialectica* **8** (1954) 37-47.
[51] Will, F. L., 'Kneale's Theories of Probability and Induction', *Philosophical Review* **63** (1954) 19-42.
[52] Wisdom, John O., *Foundations of Inference in Natural Science*, Methuen and and Company, London, 1952.
[53] von Wright, G. H., 'Catnap's Theory of Probability', *Philosophical Review* **60** (1951) 362-374.
[54] von Wright, G. H., *A Treatise on Induction and Probability*, Harcourt, Brace and World, New York, 1951.