股市和经济预测的信息和可信度[1]

鲁晨光

1. 序言

人们常说，对专家的预测不能不信，也不能全信。对于股市涨跌和经济指标的预测是这样，对于医学检验、天气预报和 GPS 等也是这样。我们应该在多大程度上相信一个预测或假设？归纳逻辑[2]研究者用信任度(degree of belief)表示信任的程度。一般说来，我们对股评家(的预测)信任度较低，然后给予信任度从低到高大概是：经济学家，天气预报，医生诊断，医学检验，GPS,手表。哲学界称：用归纳方法，通过证据或样本系列得到的信任度叫可信度或确证度(degree of confirmation)。为了符合中文习惯，本文用**信任度**表示主观的信任度(Degree of Belief)，而用客观的样本支持的信任度——即 Degree of Confirmation (DOC)，后面也用 DOC——表示**可信度**。

科学真理能通过归纳得到吗？如何评价一个科学假设？如何得到可信度？——归纳问题在当代已经演变为可信度问题。西方科学哲学界为此争论了约八十年。最早发起争论的是亨普尔[3]，卡尔纳普[4]，和波普尔等人[5]。

亨普尔提出著名的乌鸦悖论。按照数理逻辑，假设 h_1=“若 s_1 则 s_2” 和 h_2=“非 s_2 则非 s_1”等价，即“所有乌鸦是黑的”和“不是黑的就不是乌鸦”等价，所以支持 h_2 的证据也支持 h_1。据此，一根白粉笔支持“不是黑的就不是乌鸦”，从而也支持“所有乌鸦是黑的”。而根据常识，白粉笔和“所有乌鸦是黑的”不相关。所以存在悖论。

卡尔纳普(Carnap)和波普尔(Popper)则带头掀起了著名的的归纳和证伪之争。卡尔纳普认为归纳可以提高逻辑概率，可用条件逻辑概率表示可信度。而波普尔认为，用条件逻辑概率解释归纳，那还是演绎方法。科学理论只能通过假设和证伪(或检验)得到，证伪的准则是

[1]本文完成于 2016 年 9 月。发表于本人文集《股市幸存者如是说》（升级版，清华大学出版社）中了。

[2]James Hawthorne, “Inductive Logic.”[2004/2012], In *Stanford Encyclopedia of Philosophy* ed. Edward N. Zalta. http://plato.stanford.edu/entries/logic-inductive/

[3]Carl G. Hempel. 1954. “Studies in the Logic of Confirmation.” *Mind* 54:1–26 and 97–121.

[4]Rudolf Carnap. 1952. *The Continuum of Inductive Methods*. Chicago: University of Chicago Press.

[5] 波普尔，科学发现的逻辑(1934 年德文版，2008 年中译版，中国美术学院出版社

信息(语义信息)[6]——先验逻辑概率越小，如果经得起检验，信息量就越大。波普尔提出这一观点可是在仙农之前，很不简单。波普尔后来也提出过相关公式，但是没有引起足够关注(后面再讲)。倒是卡尔纳普于仙农理论(1948 年)问世后，于 1952 年和 Bar-hillel 联名发表了一篇语义信息论文章[7]，其核心公式恰恰体现了波普尔思想：信息=log(1/逻辑概率)，但是这一公式却带来另一个悖论(后面再讲)。

最近十几年，西方哲学界兴起了贝叶斯确证——用贝叶斯公式得到可信度。流行的可信度公式有 7-8 种[8](包括不流行的更多)，有人看重概率预测或似然度增量(似然度学派)，有人看重逻辑概率增量(贝叶斯学派)，谁也说服不了谁。其实两派都用贝叶斯公式，但是似然度学派用的概率可解释为频率，而贝叶斯学派用的概率是逻辑概率。他们通常用 h 表示一个假设，用 e 表示支持的证据，$\neg h$ 表示 h 的否定，$\neg e$ 表示否定 h 的证据。用 h 的概率或条件概率构成可信度，比如一个流行的确证测度是 $d(e，h)=Pr(h|e)-Pr(h)$, 就属于贝叶斯学派; 用 e 的概率或条概率构造可信度，比如 $n(e, h)=P(e|h)-P(e|\neg h)$, 就属于似然度学派。

医学界也面临检验的可信度问题。但是医学界不管哲学界的可信度测度，而是使用自己的似然度比率（简称似然比， 记为 LR）公式：阳性“+”的似然比 LR^+=真阳性比例/假阳性比例。看来它属于似然度学派。但是理想的可信度应该在 1 和-1 之间，而 LR^+最大值一般超过 1，所以医学界声称 LR^+表示检验有多可靠，言下之意是不管 LR^+是不是可信度。

可信度 $d(e, h)$比较著名，有人把它归功于卡尔纳普，有人把它归功于后来某人。它可以保证上界是 1，下界是-1。负的可信度可以表示我们对谎言的信任程度。不过有人批评说，一个正例就能得到条件逻辑概率(另外的先验逻辑概率不需要证据)，难道由一个证据就能得到可信度？辩护者说：我们计算的是可信度的增量，一个证据也能产生增量。但是，我要说，这个增量也太大了吧？假设检验阳性=h_1=“此人有病”，那么 h_1 的先验逻辑概率就是有病的比例，比如说 0.1，如果受检验者真的有病，后验逻辑概率应该是 1。于是得到 $d(e，h)=1-0.1=0.9$。如果有两个、十个证据呢？如果第十个是反例呢？可见，d 测度以及类似的测度都没有像医学界的 LR^+那样用到统计结果，都是有问题的。可信度应该来自正反例统计，个别例子只能对统计结果产生很小的影响，这是毫无疑问的！

---

[6]Karl Popper, *Conjectures and Refutations*. 1963; Repr, London and New York: Routledge.2005. p. 294.

[7]Yehoshua Bar-Hillel and Rudolf Carnap. 1952. “An Outline of a Theory of Semantic Information.” *Tech. Rep*. No. 247, Research Lab. of Electronics, MIT.

[8] Katya Tentori and others, Comparison of confirmation measures.[J]. Cognition 103 (2007)107–119

下面我们讨论如何得到上下界是 1 和-1 的，基于正反例统计的可信度。这要从 Bar-hillel 和卡尔纳普的那个信息公式引起的悖论 BCP(Bar-hillel-CarnapParadox)谈起。

2. 从 Bar-hillel-Carnap 悖论到可信度

有人会说："你写错了吧？应该是乌鸦悖论。"没错，就是 Bar-hillel-Carnap 悖论，它就是信息公式 I=log(1/逻辑概率)带来的，它关系到归纳逻辑的核心问题——可信度问题。

按照这一公式，逻辑概率越小，信息量就越大。那么矛盾句的逻辑概率小——为 0。由公式算出的信息量是 log(1/0)=无穷大。假话错话信息量多，真话信息量少，这违反常识。所以 Floridi 称之为 Bar-hillel-Carnap 悖论(简记 BCP)[9]。Floridi 仿照波普尔，强调信息需要似真度(truthlikeness or verisimilitude)，为此提出一套复杂公式(网上可见[10])。

我比 Floridi 早十年就提出一个简单的、考虑了似真度的信息公式[11]，1999 年发表了英文长文介绍了我的广义信息理论[12]。可惜标题没有"语义信息"，期刊名字好像与信息及哲学无关，所以没有引起足够关注。下面以股市预测为例简要介绍我的语义信息公式，以及如何用它解决 BCP。

我们用用小写字母 $e_1, e_2,\ldots$，m_e 表示具体股市指数，用大写字母 E 表示指数变量，它是集合 A 中的一个元素，即 $E\in A=\{e_1, e_2, \ldots, e_m\}$。$E=e_i$ 表示 e_i 发生。类似地，一个假设或预测 H 是一个变量，$H\in B=\{h_1, h_2, \ldots, h_n\}$。$H=h_j$ 表示 h_j 被选择。一个预测 h_j 发生后，e_i 出现，即 $E=e_i$，那么就有命题 $h_j(e_i)$及其真假。我们记命题 $h_j(e_i)$的真值为 $T(h_j|e_i)$，记谓词 $h_j(e_i)$的真值函数为 $T(h_j|E)$。记 A_j 是使 h_j 为真的元素构成的集合，如果 $e_i\in A_j$, $T(h_j|e_i)=1$；否则 $T(h_j|e_i)=0$.

[9]Luciano Floridi. 2004. "Outline of a theory of strongly semantic information." *Minds and Machines* 14:197-221.

[10]Luciano Floridi. "Semantic conceptions of information." in *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. http://plato.stanford.edu/entries/information-semantic/

[11]鲁晨光，《广义信息论》，[M]. 1993, 中国科技大学出版社。

[12]鲁晨光，A generalization of Shannon's information theory." *Int. J. of General Systems,* 1999，28 (6): 453-490.

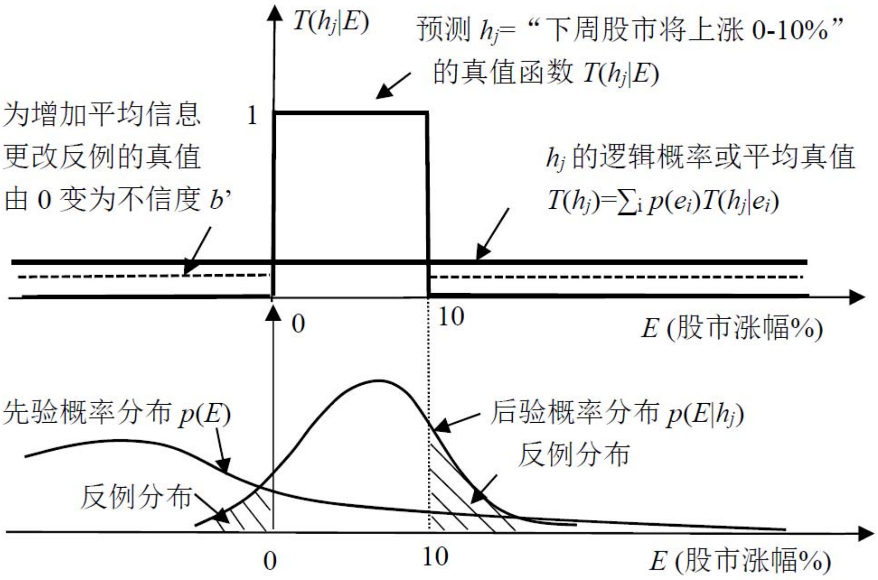


图 1. 股市预测的真值函数和反例的真值 b’的调整

指数 E 从 0 到 10%构成使预测 h_j 为真的集合 A_j=[0,10], E 超出这个范围就是预测的反例。图中我们且假设真值函数是经典的，即仅有 0 和 1 二值。实际上，我们不能说涨 9.9%或 0.1%预测就 100%正确，而跌 0.1%或涨 11%就是 100%错误。人脑实际用的真值函数是模糊的，即是一条山形曲线，而不是一条直上直下的矩形线。但是为了解释方便，我们先用图 1 说明正反例并简化信息计算，后面我们再改用山形真值函数说明实际的语义信息。

我的语义信息量公式是：

$$I(e_i;h_j)=\log\frac{T(h_j|e_i)}{T(h_j)} \qquad (1)$$

其中 $T(h_j)$是平均真值或逻辑概率是：

$$T(h_j)=\sum_i P(e_i)T(h_j\,|\,e_i) \qquad (2)$$

如果预测后指数 E 确实在 E 在 A_j 中，$T(h_j|e_i)$就等于 1，上面公式就变成 $I(e_i;h_j)=\log[1/T(h_j)]$, 类似于 Barhillel 和卡尔纳普的信息公式。但是加上分子 $T(h_j|e_i)$就可以避免 BCP。对于矛盾句，$T(h_j|E)\equiv0$ 且 $T(h_j)=0$，所以 $I(e_i;h_j)=\log(0/0)$。因为 $\log(0_+/0_+)=\log1=0$ (0_+是正无穷小)。所以认为对于矛盾句 $I(e_i;h_j)=0$ 是合理的。上面信息公式也可以保证永真句信息是 0，如波普尔所肯定。对于永真句，$T(h_j|E)\equiv1$, $T(h_j)=p(e_1)+p(e_2)+\ldots+p(e_m)=1$; 所以 $I(e_i;h_j)=\log(1/1)=0$。

根据推广的贝叶斯公式，即语义贝叶斯公式[13]，假设 h_j 真，预测的 E 发生的条件概率就是：

$$p(E \mid h_j\text{是真的}) = \frac{p(E)T(h_j \mid E)}{T(h_j)} \qquad (3)$$

一般情况下，$p(E|h_j$ 是真的$)\neq p(E|h_j)$。$p(E|h_j)$是样本概率分布，是用以检验假设的，反例的概率不一定是 0。而 $p(E|h_j$ 是真的$)$是理论预测的概率，反例的概率为 0。下面我们简单地把 $p(E|h_j$ 是真的$)$简记为 $p(E|A_j)$。根据式(1)和(3)，我们得到：

$$I(e_i;h_j) = \log\frac{p(e_i \mid A_j)}{p(e_i)} \qquad (4)$$

这一公式类似于波普尔的检验严厉性公式[14]和 Milne 提出的可信度公式[15]。不同的是，公式(4)中，1)h_j 变成 A_j, 它表明：h_j *是真的*，而不是：h_j *被选择*; 2)省略了背景知识 b。本文中 $P(E)$已经包含了背景知识，如果没有背景知识，E 就是等概率的，即 $P(E)\equiv$常数。式(3)和(4)和经典的相对信息公式类似，从而兼容经典信息论。

图 1 中 $T(h_j)$大约是 1/8, 当指数涨幅 E 大于 0 且小于 10 时，信息是 log[1/(1/8)]=3 比特。当小于 0 或大于 10%，信息是负无穷大(-∞）。对 $I(e_i;h_j)$求平均，我们得到平均语义信息公式

$$I(E;h_j) = \sum_i p(e_i \mid h_j)\log\frac{T(h_j \mid e_i)}{T(h_j)} \qquad (5)$$

和广义 Kullback-Leibler 公式(是仙农互信息公式在 H=h_j 时的特例)：

$$I(E;h_j) = \sum_i p(e_i \mid h_j)\log\frac{p(e_i|A_j)}{p(e_i)} \qquad (6)$$

当有一个反例存在，上面平均信息还是负无穷大。平均信息负无穷大正好反映波普尔的证伪思想——一个反例就足以证伪一个全称假设，不管正例有多少。但是负无穷大信息不符合常识。难道股市预测正确 100 次的信息增量还抵消不了错误一次的信息损失？为此，我以前用真值函数在 0 和 1 之间变化的模糊假设取代真值仅取 0,1 二值的经典假设。现在我发现可以把反例的真值从 0 调整到 *b*’——表示我们对假设的不信度(用 b 表示信任度)，这样也可

---

[13] 鲁晨光，广义信息论，中国科学技术大学出版社， 1993 年。p. 37.

[14]Karl Popper, *Conjectures and Refutations*. 1963; Repr 2005. London and New York: Routledge. p. 526.

[15] Peter Milne. 1996. “log[P(h/eb)/P(h/b)] Is the One True Measure of Confirmation.” *Philosophy of Science* 63:21-26.

以避免负无穷大信息。适当调整 b'可以使平均语义信息达最大，等于 Kullback-Leibler 信息。有了优化的不信度，记为 b'^*，我们就可以得到优化的信任度，即可信度(或确证度)b^*。

下一节谈不信度优化。下面先谈为什么最优真值函数往往是模糊的。式(6)可以写成两个 Kullback-Leibler 信息差的形式：

$$I(E;h_j)=\sum_i p(e_i\mid h_j)\log\frac{p(e_i|h_j)}{p(e_i)}-\sum_i p(e_i\mid h_j)\log\frac{p(e_i|h_j)}{p(e_i\mid A_j)} \qquad (7)$$

因为 Kullback-Leibler 信息总是大于 0，所以上式第二项为 0 的时候，平均信息量最大。容易证明，当理论预测和样本统计符合的时候，即 $p(E|A_j)=p(E|h_j)$，或 $T(h_j|E)$正比于 $P(h_j|E)$时，第二项等于 0，这时平均信息量达最大。令 $T(h_j|E)$的最大值是 1，可以得到

$$T(h_j|E)=P(h_j|E)/P(h_j|e_j^*)=P(E|h_j)/P(E)/[p(e_j^*|h_j)/p(e_j^*)] \qquad (8)$$

其中 e_j^*是使 $P(h_j|E)$达最大的 E。上式似乎反映了贝叶斯学派和似然度学派的统一。不过上式中没有逻辑概率，也没有预测的概率，只有客观统计得到的概率。条件概率 $P(h_j|E)$可被称之为选择概率，反映预测者选择语句的规律，一般是一条曲线，所以最好的真值函数也应是一条曲线。这就是为什么听众对天气预报和股市预测一般按模糊方式理解。

上面所有公式对于模糊真值函数也是合适的。特别是对于数字预测或估计 h_j= “E 大约是 e_j”，比如 h_j= “指数大约会涨到 10000 点”,真值函数可假设为 $T(h_j|E)=\exp[-(E-e_j)^2/(2d^2)]$(其中 d 是标准偏差)，其信息量随 E 变化如图 2 所示。

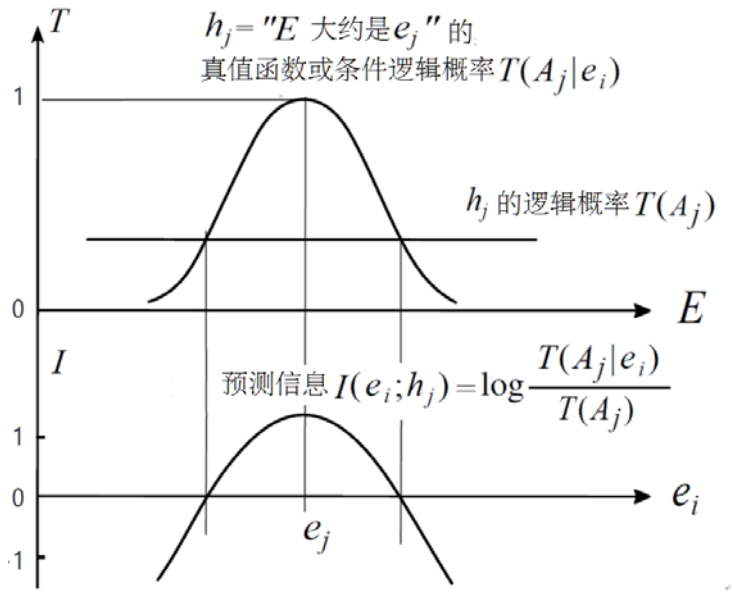


图 2.预测信息随偏差增大而减小，随逻辑概率减小而增大

温度表，GPS 和手表读数或指针都可以看做估计，也都可以用上述公式计算信息。什么样的预测信息量大？1)误差要小；2)逻辑概率要小。什么样的预测逻辑概率小？根据公式

(2)，首先，外延小的预测逻辑概率小，比如“股市涨幅大概是 10%”就比“股市会涨”逻辑概率小。第二，事件要出乎意外，比如“指数会有翻倍涨幅”就比“指数会涨 10%”逻辑概率小。股市常见的模糊假设比如:“指数可能会涨”，“如果没有利空，指数会持续缓慢上涨。”这些预测逻辑概率大，信息极少。

当然，逻辑概率小的预测，比如“指数今年会涨一倍”，也容易错，或者说容易被证伪。对于预测者来说，如果你预测不准，你用的语言就要更加模糊些。一种方法是用“...大约会涨到…”这样的语言，等于用标准差 d 较大的真值函数。第二种方法是降低听众信任度，比如说“指数有涨一倍的可能”，言下之意是不排除其他可能。下面我们就讨论如何通过降低信任度 b 提高平均语义信息，并得到优化的信任度——即可信度 b^*。

3. 优化不信度得到可信度——以股市涨跌和医学检验为例

为简化起见，我们把 E 分为两类：e_1 和 e_0, e_1 使 h_1 为真，e_0 使假设 h_1 为假。我们简记 $P_1=P(e_1)$, $P_0=P(e_0)$, $Q_1=P(e_1|h_1)$，$Q_0=P(e_0|h_1)$。

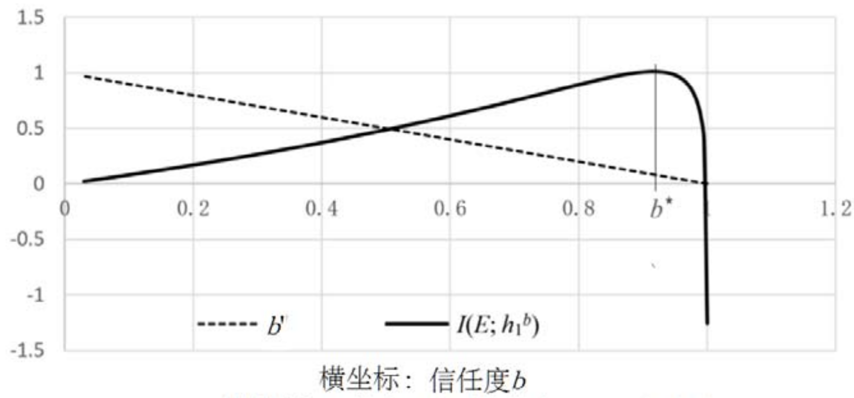
我们用 h_j^b 表示可信度为 b 的 h_j(注意 $h_j^1=h_j$)，其真值函数定义为(假设 b>0)：

$$T(h_j^b|E)= b' +bT(h_j|E) \qquad (9)$$

其中 b'是不信度，$b'=1-|b|$，它表示一个假设中含有永真句的比例。那么就有 $T(h_j^b|e_1)=1$, $T(h_j^b|e_0)= b'$。根据式(2)，$T(h_1^b)=b'P_0+P_1$。再根据式(5)，有

$$I(E;h_1^b)=Q_0\log\frac{b'}{b'P_0+P_1}+Q_1\log\frac{1}{b'P_0+P_1} \qquad (10)$$

可以验算，b'=0 时(表示我们对 h_1 绝对相信)，有一个反例存在，平均信息就是-∞(参看图 2)。当 b'=1 时，表示假设变成永真句，信息是 0。$I(E; h_1^b)$随 b'变化见图 2。现在我们寻求使 $I(E; h_1^b)$达最大的 b'，或最优的 b^*。



横坐标：信任度b
纵坐标：不信度b'和平均信息 $I(E; h_1^b)$(比特)

图 3. 平均信息 $I(E; h_1^b)$随不信度 b'变化(设 P_0/P_1=0.8/0.2; Q_0/Q_1=0.25/0.75)

令 $I(E; h_1^b)/db'=0$，我们得到 $Q_0P_1-Q_1P_0b'=0$。解得 $b'=(Q_0/Q_1)/(P_0/P_1)$。因为 $b'=(Q_0/Q_1)/(P_0/P_1)$时, $I(E; h_1^b)$相对 b'的二阶导数小于 0，所以 $b'=(Q_0/Q_1)/(P_0/P_1)$就是最优的 b'，记为 b'^*, 它使信息 $I(E; h_1^b)$达到其最大值 $I(E; h_1^{b^*})$。

因为根据公式(8)，有

$$b'^*=(Q_0/Q_1)/(P_0/P_1)=P(h_1|e_0)/P(h_1|e_1) \qquad (11)$$

根据式(9), 优化的可信度是

$$b^*=1-b'^*=1-(Q_0/Q_1)/(P_0/P_1)=1-P(h_1|e_0)/P(h_1|e_1) \qquad (12)$$

和流行的可信度公式不同之处是：1)它用样本的统计概率分布而不是单个样本计算可信度；2)它只包含客观的统计概率，并不包含预测的概率，更不包含逻辑概率；3)更加重视反例的统计概率。

对于图 1 所示股市预测，假设正例的先验概率是 P_1=0.2，有预报后，正例的后验概率 Q_1 是 0.75，那么就有

b'^*=(0.25/0.75)/(0.8/0.2)=0.083；

$b^*=1-b'^*$=0.917；

$T(h_1)$=0.2+0.083*0.8=0.267;

$I(E; h_1^{b^*})$=0.25log(0.083/0.267)+0.75log(1/0.267)=1.01 比特

相应的 Kullback-Leibler (KL)信息 $I_{KL}(E; h_1)$也是 1.01 比特。我们可以这样理解，KL 信息是客观信息，语义信息是主观信息。客观信息是主观信息的上限，优化信任度可以使主观信息达到其上限——客观信息。

注意，上式只合适$(Q_0/Q_1)<(P_0/P_1)$的情况，否则 h_1 的确证度是负的。对于负的可信度，我们定义其真值函数为(假设 $b<0$)：

$$T(h_j^b|E) =1+bT(h_j|E) \qquad (13)$$

这样定义可以保证，肯定句 h_1 有负的信任度 b 等价于其否定句 h_0 有正的信任度$|b|$，即 $b<0$ 时，$T(h_1^b|E)=T(h_0^{|b|}|E)$。类似也有 $b<0$ 时，$T(h_0^b|E)=T(h_1^{|b|}|E)$，即有负的信任度的否定句等价于有正的信任度肯定句。可以证明最优负的信任度是

$$b^*= (P_0/P_1)/(Q_0/Q_1)-1=P(h_1|e_1)/P(h_1|e_0)-1 \qquad (14)$$

假如股市预测只有两种：预测涨和预测跌，那么股市预测的信息及可信度就和医学检验

阳性和阴性类似。下面说明如何仿照医学检验方法推导出仙农信道，即 $P(H|E)$, 进而计算股市预测的可信度。正反例符号及相应个数的符号见表 1.

表 1 两种推理以及它们的正例和反例数目

	真涨或真有病 e_1	真跌或真没病 e_0
h_1=“预测涨”或阳性+	n_{11} 个正例 e_{11}	n_{10} 个反例 e_{10}
h_0=“预测跌”或阴性-	n_{01} 个反例 e_{01}	n_{00} 个正例 e_{00}

医学界把有病者 e_1 被检验出阳性(真阳性)的比例叫做敏感性(sensibility)。它就是条件概率 $P(h_1|e_1)=n_{11}/(n_{11}+n_{01})$ (假设病例很多)；把没病者 e_0 检验出阴性(真阴性)的比例叫做特异性(specificity),它就是条件概率 $P(h_0|e_1)=n_{00}/(n_{00}+n_{10})$。那么就有仙农信道如表 2 所示。

表 2 医学检验的敏感性和特异性构成仙农信道 $P(H|E)$

	真涨或真有病 e_1	真跌或真没病 e_0
$P(h_1|E)$	敏感性=$P(h_1|e_1)$	1-特异性=1-$P(h_0|e_0)$=$P(h_1|e_0)$
$P(h_0|E)$	1-敏感性=1-$P(h_1|\mathrm{e}_1)$=$P(h_0|e_1)$	特异性=$P(h_0|e_0)$

根据式(12)，h_1 或阳性的可信度就是

$$b^*=1-P(h_1|e_0)/P(h_1|e_1)=1-n_{10}/(n_{10}+n_{00})/[n_{11}/(n_{11}+n_{01})] \qquad (13)$$

阴性的可信度是：

$$b_0^*=1-P(h_0|e_1)]/P(h_0|e_0)=1-n_{01}/(n_{11}+n_{01})/[n_{00}/(n_{00}+n_{10})] \qquad (14)$$

敏感性和特异性构成了仙农信道 $P(H|E)$，确证实际上是用仙农信道确证语义信道 $T(H|E)$——见表 3。给假设或预测以合适的信任度，语义信道 $T(H|E)$就和仙农信道 $P(H|E)$ 符合或成正比，h_1(或 H)的平均语义信息就达到其上限——Kullback-Leibler 信息(或仙农互信息)。

表 3 两个不信度构成和仙农信道匹配的语义信道

	真涨或真有病 e_1	真跌或真没病 e_0
$T(h_1^{b^*}|E)$	1	$b'^*=P(h_1|e_0)/P(h_1|e_1)$
$T(h_0^{b_0*}|E)$	$b_0'^*=P(h_0|e_1)/P(h_0|e_0)$	1

因为医学上常用的阳性似然度比率 LR^+=敏感性/(1-特异性)=$1/b'^*$，所以 $b^*=1-1/\mathrm{LR}^+$。可

见上述可信度 b^*和 LR^+有正相关关系。但是 b^*的最大值是 1。如果有人搞错了，说阳性表示没病，正反例颠倒，则可以算出 b^*是负的。所以 b^*在 1 和-1 之间。

用语义贝叶斯公式，我们可以用否证度 b'^*和先验概率 $P(e_1)$计算 h_1 预测的 e_1 发生的概率(似然度)，记为 $P(e_1|h_1^{b^*})$：

$$P(e_1 \mid h_1^{b^*}) = P(e_1)/[P(e_1) + b_1'^* P(e_0)] \qquad (15)$$

用这个公式计算正确率和用贝叶斯公式计算的结果一样，但是更加方便。这个公式对于 $P(E)$变化时，预测疾病和股市上涨的概率将非常有用。比如，张三预测股市上涨的先验概率是 0.6，听到股评家李四预测上涨后——假设张三对李四预测上涨的可信度(来自过去统计)是 0.8，那么张三预测股市上涨的后验概率就是：$P(e_1|h_1^{b^*})$=0.6/(0.6+0.2*0.4)=0.6/0.68=0.88。

请注意，上述可信度更加重视反例——肯定较少的反例比较多的正例更重要。具体说来就是：阳性的可信度主要取决于无病被检验为阴性的正确率；阴性的可信度主要取决于有病被检验为阳性的正确率。类似地，一个股评家预测股市上涨的可信度主要取决于他以前预测股市下跌的的正确率；预测股市下跌的可信度主要取决于他过去预测上涨的正确率。

这个结论和常识似乎不同，但是也可以理解。按照这个公式，过去预测房价上涨总是对的任志强如果预测房价下跌，那么他的预测可信度就高过易宪容和谢国忠的预测（他俩过去预测房价下跌总是失败）。**其实**医学界早就知道：灵敏度可以作为避免伪阴性的量化指标，而特异度可以作为避免伪阳性的量化指标[16]。上述可信度和医学界共识是一致的。

如果除了上涨和下跌，还有持平(比如涨跌小于 3%算持平)，记为 e_2，一共三种结果。式(13)和(14)同样。但是表 2 中，“$1-P(h_0|e_0)=P(h_1|e_0)$”变成“$1-P(h_0|e_0)-P(h_0|e_2)=P(h_1|e_0)$”。

现在看乌鸦悖论就清楚了。首先，数理逻辑中，因为命题真值只有 0 和 1 二值，h_1=“若 s_1 则 s_2”和 h_2=“非 s_2 则非 s_1”等价，但是在归纳逻辑中，两者并不等价，因为虽然两者反例相同，但是正例不同，所以可信度也不同。h_1 的可信度见式(13)。h_2 的可信度应该是

$$b_2^*=1-n_{10}/(n_{11}+n_{10})/[n_{00}/(n_{00}+n_{01})] \qquad (16)$$

通过式(13)，可以求出不同证据 e_{00} 和 e_{11} 引起可信度 b^*的不同增量(不缀)。

第二，我们从来没有见过不黑的乌鸦，所以 e_{10} 的个数 n_{10} 是 0，h_1 的可信度就是 1，不管 e_{00} 和 e_{11} 是多少。说白粉笔不支持“所有乌鸦是黑的”也可以。但是换成“所有天鹅是白的”或“HIV 检验呈阳性的人带有艾滋病病毒”，情况就不同了。

[16]佚名，灵敏度和特异度.维基百科.[2015-4-4].

第三，n_{10}不是 0 的时候，可信度公式(13)表明，e_{00}不但支持 h_1, 而且可能比 e_{11}支持更有力——这一结论出乎大家(除了医学检验研究者)预料。因为 e_{00}比 e_{10}多则特异性大，反例比例少。从式(13)可见，当 n_{00}比 n_{10}大很多且 $n_{11}= n_{01}$时，随着 n_{00}加倍，b'^*差不多减少一半。然而，当 n_{11}加倍的时候，$b_1'^*$只减少 1/4。

我们可以把医学上的敏感性和特异性概念推广到股市预测，那就是令：敏感性=上涨预测成功率；特异性=下跌预测成功率。于是有：预测上涨的可信度=1-(1-特异性)/敏感性；预测下跌的可信度=1-(1-敏感性)/特异性。

4. 预测股市指数或经济指标的可信度

现在考虑“指数涨幅大约是 10%”这样的预测。其一般形式是 h_j=“$E\approx e_j$”。经济指标预测和各种数量预测同理。甚至温度表读数、GPS 定位，都可以看做这样的预测。这类预测数学上叫估计，用 $\hat{e}_j$表示，即 $h_j=\hat{e}_j$=“$E\approx e_j$”。

我们用 Z 表示观察条件 $z_1, z_2,\cdots, z_w$中的一个，$z_1, z_2,\cdots, z_w$构成集合 C，C 中有子集 C_j, 当 Z 在集合条件 C_j中，就有预测 h_j, 这样就有样本的条件概率 $P(E|h_j)=P(E|Z\in C_j)$, 简记为 $P(E|C_j)$。根据 $P(E|C_j)$和先验概率 $P(E)$可以得到 h_j的预报规则函数

$$P(h_j|E)=P(C_j|E)=KP(E|C_j)/P(E) \qquad (17)$$

其中 K 是常数。$P(h_j|E)$通常接近于正态分布(比如 GPS 箭头相对实际位置就接近于正态分布), 如果预测较大偏差也是可能的，则假设它加上噪声分布 c。再假设预测含有系统偏差 Δe，那么就有

$$P(h_j|E)=K\exp[-(E-\Delta e-e_j)^2/(2d^2)]+c \qquad (18)$$

h_j本身就是模糊的，c 则决定了它的可信度，在可信度小于 1 时，预测就更加模糊。设优化的预测的真值函数(j 变成 k)为

$$T(h_k^{b^*}|E)=b^*\exp[-(E-e_k)^2/(2d^{*2})]+b'^* \qquad (19)$$

通过式(8)可以推导出，$\hat{e}_k=\hat{e}_j+\Delta e$, $d^*=d$, $b^*=1-c/(K+c)$时，平均语义信息 $I(E; h_k^{b^*})$达最大。

$b^*=1-c/(K+c)$表明，噪声比例越小，预测的可信度越高。

如果函数$P(h_j|E)$的形式是不知道的，或者说我们找不到使$T(h_k^{b*}|E)$正比于$P(h_j|E)$的函数，则h_j的可信度是使平均语义信息量达最大的信任度b, 即

$$b^* = \arg\max_b I(E;h_j^b) = \arg\max_b \sum_i P(e_i|h_j)\log\frac{T(h_j^b|e_i)}{T(h_j^b)} \qquad (20)$$

这也是可信度的一般定义。

关于由来已久的证伪和确证之争，现在可以得到结论：1)证伪(包括假设的检验、选择和优化)用语义信息准则——这意味着，一个假设越正确，精度越高，先验越是意外，经确证后越是可信，则它提供的信息就越多，就越是可以接受； 2)确证就是通过反例和正例的相对比例优化假设的信任度得到可信度，从而提高平均语义信息，因此是证伪的辅助手段。

5. 总结

平均语义信息公式是本文用来求解可信度的主要工具，它和仙农信息论及波普尔理论兼容。新的语义信息公式在本质上是用客观的样本的概率分布 $P(E|h_j)$检验主观预测的概率分布$P(E|A_j)$从而得到平均信息量；而新的确证方法在本质上是用客观的仙农信道($P(h_j|E)$, j=1, 2,…)确证主观的语义信道(即真值函数 $T(h_j|E)$, j=1, 2, …)。新的方法可以同时解决Bar-hilel-Carnap悖论和乌鸦悖论。其实际应用的合理性可以通过股市预测和医学检验看出。