定量预测方法有效性指标分析*

李文华
(郑州工业大学数力系)

摘　要　定义了连续情形的预测精度公式，在此基础上给出了1套衡量定量预测方法有效性指标的一般定义，并讨论了几类有效度之间的关系，同时给出了估计拟合有效度、预测有效度的具体方法。从而给预测工作者选择有效的预测方法提供了可靠的理论基础，减少预测工作的盲目性和预测费用。

关键词　预测；精度；有效度

中图分类号　F222

在进行预测时，预测方法是否有效，直接影响着预测的质量和决策的科学性。因此，选择有效的预测方法具有重要意义。

文献[1]以文献[2]为基础，提出了用精度描述预测方法有效性的有效度指标，并引入了拟合有效度、预测有效度、综合有效度等概念。但并没有给出预测有效度指标的算法，因而很难在预测之前判定预测方法的有效性，同时它也没有考虑各点的精度加权问题，从而使有效度的应用受到很大限制。

鉴于以上分析，本文仍用文献[1,2]的有关名词术语，给出新的有效度定义，并给出估计方法。

1　精度公式的讨论

设$[0, T_0]$为样本区间，$[T_0, T_0+T]$为预测区间，$y=g(t)$为真值函数，$y=h(t)$为某种预测函数，且$h(t)$，$g(t)$在$[0, T_0+T]$上 L 可积，$g(t)$在$[0, T_0]$有最大值 $M=\max\limits_{[0, T_0]}|g(t)|$。

设$A(t)$表示精度，令$B(t)=1-A(t)$，则$B(t)$表示t时刻预测值与实际值之间的相对误差。定义精度公式为：

$$A_1(t)=\begin{cases} 1-\dfrac{|h(t)-g(t)|}{M} & |h(t)-g(t)|\leqslant M \quad \text{ao} \\ 0 & |h(t)-g(t)|>M \end{cases}$$

$$A_2(t)=\begin{cases} 1-\dfrac{|h(t)-g(t)|}{|g(t)|} & |h(t)-g(t)|\leqslant|g(t)| \\ 0 & |h(t)-g(t)|>|g(t)| \end{cases}$$

那么

$$B_1(t)=\begin{cases} \dfrac{|h(t)-g(t)|}{M} & |h(t)-g(t)|\leqslant M \quad \text{at} \\ 1 & |h(t)-g(t)|>M \end{cases}$$

*河南省科委自然科学基金资助项目(954051600)
收稿日期：1997－12－15

第一作者　女　1965年1月生　硕士学位　讲师

$$B_2(t)=\begin{cases}\dfrac{|h(t)-g(t)|}{|g(t)|} & |h(t)-g(t)|\leqslant|g(t)| \\ 1 & |h(t)-g(t)|>|g(t)|\end{cases}$$
衡量

说明：

(1)上式中$0\leqslant A(t)\leqslant 1$，它体现了时刻$t$处预测值$h(t)$与其真值$g(t)$的接近程度。$A(t)$越大，$h(t)$越优；$A(t)$越小，$h(t)$越劣。同样，$B(t)$越小，$h(t)$越优；$B(t)$越大，$h(t)$越劣。

(2)若$|g(t)|$在$[0, T_0+T]$无可靠的较大的正值下界，如$g(0)>0$且$g(t)$递减，$g(t)$有可能在$[0, T_0+T]$递减到0，则$A_2(t)$有可能失去可比性，此时宜用$A_1(t)$。

此定义的优点是，它不仅包含了其它文献的精度定义，而且因$h(t)$，$g(t)$是连续函数，从而可利用积分给出更一般的有效度的定义。

2　预测方法有效度的定义

定义1　若函数$\varphi_s(t)\geqslant 0$，$t\in[0, T_0]$，$\int_0^{T_0}\varphi_s(t)\mathrm{d}t=1$，则称$m_s(h,g)=1-\left[\int_0^{T_0}\varphi_s(t)\,|B(t)|^p dt\right]^{\frac{1}{p}}$为模型$y=h(t)$对$y=g(t)$的拟合有效度。

其中$\varphi_s(t)$称为侧重性权函数，$0<p<+\infty$为参数。

定义2　若函数$\varphi_f(t)\geqslant 0$，$t\in[T_0, T_0+T]$，$\int_{T_0}^{T_0+T}\varphi_f(t)\mathrm{d}t=1$，则称$m_f(h,g)=1-\left[\int_{T_0}^{T_0+T}\varphi_f(t)\,|B(t)|^p dt\right]^{\frac{1}{p}}$为模型$y=h(t)$对$y=g(t)$的预测有效度。，称$\overline{m}_f(h,g)=1-m_f(h,g)$为预测无效度。

其中$\varphi_f(t)$称为侧重性权函数，$0<p<+\infty$为参数。

定义3　若$\varphi(t)=\begin{cases}\alpha\varphi_s(t) & t\in[0, T_0] \\ (1-\alpha)\varphi_f(t) & t\in(T_0, T_0+T]\end{cases}$　$(0<\alpha\leqslant 1)$，

则有$\int_0^{T_0+T}\varphi(t)\mathrm{d}t=1$，称$m(h,g)=1-\left[\int_0^{T_0+T}\varphi(t)\,|B(t)|^p dt\right]^{\frac{1}{p}}$为模型$y=h(t)$对$y=g(t)$的综合有效度，称$\overline{m}(h,g)=1-m(h,g)$为综合预测无效度。其中$\alpha$越大，越侧重拟合有效度，特别地$\alpha=1$时，$m(h,g)=m_s(h,g)$。

在定义1,2,3中，常用的定义形式为$p=1$和$p=2$的情形。当$p=1$时，m，m_s，m_f可认为是加权平均精度；$p=2$时，m，m_s，m_f可认为是加权均方精度，有较强的直观含义和实际含义。$\varphi_s(t)$，$\varphi_f(t)$，$\varphi(t)$的取法要根据具体情况作适当选取。

3　关于几类有效度关系的两个定理

定理1　$m(h,g)$，$m_s(h,g)$，$m_f(h,g)$有如下关系：

$$(1-m)^p=\alpha(1-m_s)^p+(1-\alpha)(1-m_f)^p\text{，即 }\overline{m}^p=\alpha\overline{m}_s^{\,p}+(1-\alpha)\overline{m}_f^{\,p}$$

证明：由定义

$$(1-m)^p=\int_0^{T_0+T}\varphi(t)\mid B(t)\mid^p dt$$
$$=\int_0^{T_0}\alpha\varphi_s(t)\mid B(t)\mid^p dt+\int_{T_0}^{T_0+T}(1-\alpha)\varphi_f(t)\mid B(t)\mid^p dt$$
$$=\alpha(1-m_s)^p+(1-\alpha)(1-m_f)^p$$

证毕。

定理 2 若侧重性权函数为等权情形，即当 $\varphi_s(t)=\frac{1}{T_0}$，$\varphi_f(t)=\frac{1}{T}$，$\varphi(t)=\frac{1}{T_0+T}$时，则有 $\alpha=\frac{T_0}{T+T_0}$.

证明： $\because$ $\varphi(t)=\begin{cases}\alpha\varphi_s(t) & t\in[0,T_0]\\(1-\alpha)\varphi_f(t) & t\in[T_0,T_0+T]\end{cases}$ $\therefore$ 当 $t\in[0,T_0]$时，有 $\frac{1}{T_0+T}=\alpha\frac{1}{T_0}$，当 $t\in[T_0,T_0+T]$时，有$\frac{1}{T_0+T}=(1-\alpha)\frac{1}{T}$，解之：$\alpha=\frac{T_0}{T_0+T}$。

即侧重性权函数为等权时，有 $\alpha=\frac{T_0}{T_0+T}$。

从上式可以看出，预测期越短(T 越小)，专家能提供的主观信息越少，α 越大，此时预测越偏重拟合效果；预测期越长(T 越大)，历史资料相对少，主要靠专家提供主观预测，α 越小，此时偏重预测效果。非等权情形也具有此特征。

需要说明的是，尽管以上 3 个定义较完善，但通常在使用时仍不用积分形式，而是用离散的形式估计有效度的大小。

4 拟合有效度的估计

设 $y_t=g(t)$，$t=1,2,\cdots,N$ 为观测值，$h(t)$为预测模型 $y=h(t)$在观测点的取值，则 $h(t)$对 $g(t)$的拟合有效度有一般估计式：($T_0=N$)

$$\hat{m}_s(h,g)=1-\Big[\sum_{t=1}^{N}\varphi_s(t)B^p(t)\Big]^{\frac{1}{p}}$$

其中
$$B(t)=\frac{\mid h(t)-y_t\mid}{M}\text{ 或 }B(t)=\frac{\mid h(t)-y_t\mid}{\mid y_t\mid}$$
$$M=\max y_t,\ t=1,2,\cdots,N$$

在$[0,T_0]$内，当 N 个观测点的时间间隔 ΔT 的最大值 $\lambda=\max_i\{\Delta T_i\}\to 0$ 时，$\hat{m}_s(h,g)\to m_s(h,g)$。

5 预测有效度的估计(平均值法)

步骤 1 用特尔斐法确定预测值的取值区间

考虑$[T_0,T_0+T]$上的 m 个点，对此 m 个时刻向专家问卷(同时给专家提供过去的有关数据和资料)，当通过几轮的应答，专家意见日趋一致时，即得出第 i 个预测点处预测值的最高值 $\hat{y}_i^{(2)}$和最低值 $\hat{y}_i^{(1)}$($i=1,2,\cdots,m$)，也即得到每个点处的预测值区间$[\hat{y}_i^{(1)},\hat{y}_i^{(2)}]$。

步骤 2　确定拟合专家预测带

以 $\hat{y}_1{}^{(1)}$，$\hat{y}_2{}^{(1)}$，…，$\hat{y}_m{}^{(1)}$ 和 $\hat{y}_1{}^{(2)}$，$\hat{y}_2{}^{(2)}$，…，$\hat{y}_m{}^{(2)}$这两组数为基础，分别拟合出两条曲线，得到一个带形区域，称为专家预测带。如图 1 所示。在每点 i 处求平均值$\hat{y}_i=\frac{1}{2}(\hat{y}_i{}^{(1)}+\hat{y}_i{}^{(2)})$（$i=1,2,\cdots,m$），由这 m 个值$\hat{y}_i$（$i=1,2,\cdots,m$）再拟合 1 条曲线(见图 1(＊))，则该曲线完全落在预测带内。

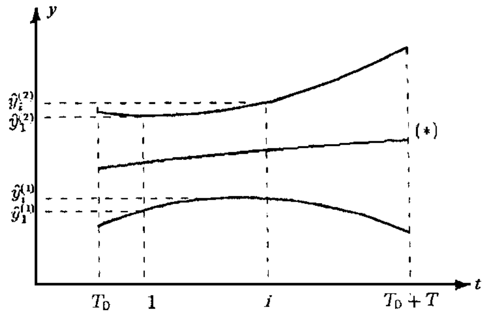

图 1　专家预测带

步骤 3　估计预测有效度

若预测方法算出的 m 个点预测值均落在预测带内，则此方法有效；至少有 1 个不落在预测带内，则预测方法无效。若落在预测带内的预测方法不止 1 个，则预测值越靠近曲线(＊)，预测方法的有效度越大。

这样，预测有效度 $=\frac{1}{m}\sum_{i=1}^{m}A_i$，其中 $A_i=1-\frac{|y_i-\hat{y}_i|}{|\hat{y}_i|}$

上式表明，预测有效度可表示为各个点的预测精度的平均值(将 $\hat{y}_i$ 视为真值)。

此方法简单易算，因依靠专家进行估计，故有一定的可靠性。

参考文献

1　王明涛.预测方法的有效性分析.预测，1994(6)：57～59
2　唐小我.简单平均预测方法的有效性分析.预测，1992(2)：57～60
3　费革胜.预测的精度分析.预测，1990(6)：20～23
4　费革胜.再论预测的精度问题.预测，1991(6)：24～29

Analysis of Availability Index of Quantitative Forecast Method

Li Wenhua
(*Zhengzhou University of Technology*)

Abstract　The formula of forecast precision under the condition of continuation is defined, and accordingly, a series of general definations of evaluating the availability index of the quantitative forecast method are given. The relations of sereral availableness are discussed and the actual method of evaluating the fitted availableness and the forecast availableness is put forward finally.

Keywords　forecast；precision；availableness