

一种带判别函数的单样本有限状态分布检验

张夏怡

华南师范大学数学科学学院, 广东 广州

收稿日期: 2022年7月18日; 录用日期: 2022年7月29日; 发布日期: 2022年8月10日

摘要

本文引进了一种带判别函数的单样本有限状态分布检验, 并给出了基于精确分布的检验计算方法。在本文方法中, 给出不同的判别函数, 检验可表示不同的具体检验方法, 其中几个经典的检验方法, 例如 X^2 拟合优度检验、二项检验和似然比检验都是本文检验方法在不同判别函数下的特例。本文给出了一个由样本的分布频率和总体的概率的差距构成的判别函数。案例分析显示, 这个检验方法比上述三种经典的检验方法都更有效率。

关键词

有限状态分布检验, 二项检验, X^2 拟合优度检验, 似然比检验

A One-Sample Finite State Distribution Test with Discriminant Function

Xiayi Zhang

School of Mathematical Sciences, South China Normal University, Guangzhou Guangdong

Received: Jul. 18th, 2022; accepted: Jul. 29th, 2022; published: Aug. 10th, 2022

Abstract

This paper introduces a one-sample finite state distribution test with a discriminant function, and gives a test calculation method based on exact distribution. In the method of this paper, given different discriminant functions, the test can represent different specific test methods, among which several classic test methods, such as X^2 goodness-of-fit test, binomial test, and likelihood ratio test are all tested in this paper. This paper presents a discriminant function composed of the difference between the distribution frequency of the sample and the probability of the population. The case study shows that this test method is more efficient than the three classical test methods mentioned above.

Keywords

Finite State Distribution Test, Binomial Test, X^2 Goodness of Fit Test, Likelihood Ratio Test

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

有限状态分布检验有三个常用方法。第一个是 X^2 拟合优度检验方法[1] [2] [3]。设一总体有 K 个状态，其中第 i 个状态出现的概率为 $p_i, i=1, 2, \dots, K$ 。对总体进行 m 次重复抽样，其中第 i 个状态出现得 n_i 次。由皮尔孙(Pearson)定理，拟合优度检验的 X^2 统计量的极限分布为自由度为 $K-q-1$ 的卡方分布，其中 q 为原总体分布要估计的参数个数。

$$X^2 = \sum_{i=1}^K \frac{(n_i - mp_i)^2}{mp_i} \xrightarrow{m \rightarrow +\infty} X^2(K-q-1) \quad (1.1)$$

记有限状态分布的假设检验为

$$H_0 : p_i = p_{i0}, i=1, 2, \dots, K, \quad H_1 : \exists i \text{ 使 } p_i \neq p_{i0} \quad (1.2)$$

给定显著水平 α ，根据皮尔孙定理(1.1)，当 m 充分大时， X^2 拟合优度检验为

$$\begin{aligned} \text{如果 } \sum_{i=1}^K \frac{(n_i - mp_{i0})^2}{mp_{i0}} > X_{\alpha}^2(K-q-1) &\Rightarrow \text{拒绝 } H_0 \\ \text{如果 } \sum_{i=1}^K \frac{(n_i - mp_{i0})^2}{mp_{i0}} \leq X_{\alpha}^2(K-q-1) &\Rightarrow \text{接受 } H_0 \end{aligned} \quad (1.3)$$

作为一个特例，当总体只有两个状态时，其分布检验就是二项检验(binomial test)。二项检验既可以用 X^2 拟合优度检验方法，也可以用二项分布进行精确分布检验[4] [5] [6]。对总体进行 m 次独立抽样，其中状态 1 和状态 2 分别抽得 n_1 次和 n_2 次，抽样分布为

$$P(n_1, n_2, p_1, p_2) \triangleq \frac{m!}{n_1! n_2!} p_1^{n_1} p_2^{n_2} \quad (1.4)$$

记假设检验为 $H_0 : p_i = p_{i0}, i=1, 2, \dots, K, \quad H_1 : \exists p_i \neq p_{i0}$ 。在这个假设条件下，给定显著水平 α ，检验公式为

$$\begin{aligned} \text{如果 } \min\{n_1, n_2\} < B_{0.5\alpha}(m, p_{10}, p_{20}) &\Rightarrow \text{拒绝 } H_0 \\ \text{如果 } \min\{n_1, n_2\} \geq B_{0.5\alpha}(m, p_{10}, p_{20}) &\Rightarrow \text{接受 } H_0 \end{aligned} \quad (1.5)$$

其中 $B_{0.5\alpha}(m, p_{10}, p_{20})$ 为 0.5α 分位数。

有限状态分布检验的另一个常用方法是利用似然比检验方法[3] [7]。在大样本的条件下，似然比可以表示为 X^2 统计量。

$$\ln \prod_{i=1}^K \left(\frac{n_i}{mp_i} \right)^{n_i} \approx \sum_{i=1}^K \frac{(n_i - mp_i)^2}{mp_i} \quad (1.6)$$

在假设检验(1.2)的条件下，给定显著水平 α ，根据皮尔孙定理(1.1)，当 m 充分大时， X^2 拟合优度检

验为

$$\begin{aligned} \text{如果 } \ln \prod_{i=1}^K \left(\frac{n_i}{mp_i} \right)^{n_i} > X_{\alpha}^2(K-q-1) &\Rightarrow \text{拒绝 } H_0 \\ \text{如果 } \ln \prod_{i=1}^K \left(\frac{n_i}{mp_i} \right)^{n_i} \leq X_{\alpha}^2(K-q-1) &\Rightarrow \text{接受 } H_0 \end{aligned} \quad (1.7)$$

过去, 由于计算机和程序都比较落后, 在不能找到精确分布时, 我们只能借助极限分布处理假设检验问题。但这种方法有两个问题。首先, 由于真实分布构造的拒绝域被用极限分布所构造的拒绝域替代。而这两种拒绝域的差异并不能在假设检验中表现出来。因此, 用极限分布构造的假设检验含有无法评估的分布误差。我们只能寄希望于样本量足够大, 两种分布误差足够小, 使拒绝域的差异可以忽略不计。

其次, 本文提出的带判别函数的方法可以有各种不同的判别函数。即使检验有极限分布, 其极限分布也可能有不同的形式。我们只能逐个判别函数进行讨论。如果用精确分布的话, 就是多项分布, 与判别函数无关。现在, 由于计算机的发展, 用精确分布处理假设检验变得越来越可行, 特别在抽样个数不是很大的时候。

本文第二节将介绍有限状态分布检验的一种带判别函数检验方法。第三节将说明几种传统的检验方法, 例如二项分布检验方法, X^2 拟合优度检验和似然比检验方法都是带判别函数检验方法的特例。最后, 本文比较几种判别函数检验方法的检验效率。

2. 一种检验方法

设一总体有 K 个状态, 其中第 i 个状态出现的概率为 $p_i, i=1, 2, \dots, K$ 。记 $P=(p_1, p_2, \dots, p_K)$ 。对总体进行 m 次重复抽样, 其中第 i 个状态出现得 n_i 次。样本空间 Ω 由如下的 K 维向量构成, $N=(n_1, n_2, \dots, n_K)$, 其中 n_i 为非负整数, 且 $\sum_{i=1}^K n_i = m$ 。样本空间 Ω 服从 K 项分布 $M_K(m, P)$, 样本点 N 的概率为[8]

$$P\{(n_1, n_2, \dots, n_K)\} = \frac{n!}{\prod_{i=1}^K n_i!} \prod_{i=1}^K p_i^{n_i} \quad (2.1)$$

现对总体的分布进行假设检验(1.2)。记 $P_0=(p_{10}, p_{20}, \dots, p_{K0})$ 。对总体进行 m 次重复抽样, 记抽到的样本为 $M=(m_1, m_2, \dots, m_K)$ 。记假设检验的判别函数为:

$$d(M, P_0) \triangleq \sum_{i=1}^K p_{i0} \left| \frac{m_i}{m} - p_{i0} \right| \quad (2.2)$$

判别函数是样本的分布频率与目标概率的平均距离。在零假设成立的条件下, 样本空间为服从 K 项分布 $M_K(m, P_0)$ 。记

$$S(c) = \left\{ (n_1, n_2, \dots, n_K) \mid (n_1, n_2, \dots, n_K) \in \Omega, \sum_{i=1}^K p_{i0} \left| \frac{n_i}{m} - p_{i0} \right| \geq c \right\} \quad (2.3)$$

$S(c)$ 为在样本空间中, 分布频率与假设概率平均距离大于等于 c 的点组成的集合。

在零假设成立的条件下, 样本的分布频率应该与假设概率相差不大。如果样本的分布频率与概率相差较大, 那说明抽样出现不正常的情况。因此, $S(c)$ 是由异常样本构成的集合。如果异常点集合的概率很小, 它就是拒绝域。其思路是, 如果样本落在拒绝域上, 它表明异常的小概率事件发生, 统计学上的“反例”出现, 推翻了原假设。

给定显著水平 $\alpha (\alpha > 0)$, 假设检验的拒绝域定义为

$$\text{Rej}(\alpha) = \bigcup_{c>0} \{S(c) | P\{S(c)\} < \alpha\} \quad (2.4)$$

简单地说, 拒绝域就是在 $\text{Rej}(c)$ 中概率小于 α 的最大集合, 而接受域就是 $\Omega \setminus \text{Rej}(c)$ 。

在检验中, 我们可以通过计算检验的概率值 p 来判断零假设是否被拒绝。根据检验的定义, 概率值表达如下,

$$p = P \left\{ (n_1, n_2, \dots, n_K) | (n_1, n_2, \dots, n_K) \in \Omega, \sum_{i=1}^K p_{i0} \left| \frac{n_i}{m} - p_{i0} \right| \geq \sum_{i=1}^K p_{i0} \left| \frac{m_i}{m} - p_{i0} \right| \right\} \quad (2.5)$$

它是由满足某个条件的样本构成的集合的概率, 这个条件是样本的分布频率与目标概率的平均距离大于等于抽到的样本的分布频率与目标概率的平均距离。

如果概率值 p 小于显著水平 α , 我们就拒绝零假设; 否则我们就接受(不拒绝)零假设。这个结论的证明很简单。从(1.6)可知, 如果概率值 p 小于显著水平 α , 说明抽到的样本落在拒绝域里, 因此拒绝零假设。如果概率值 p 大于等于显著水平 α , 说明抽到的样本落在接受域里, 因此接受零假设。

在重复抽样个数不是很大的时候, 上述概率值可利用精确分布(K 项分布)计算。

例 1: 设一色子在随机抛掷中每个数字朝上的概率相等。现随机抛掷 100 次, 数字 1 到数字 6 朝上的频数分别为 $m_1 = 12, m_2 = 25, m_3 = 13, m_4 = 12, m_5 = 23, m_6 = 15$ 。给定显著水平 $\alpha = 0.05$, 进行检验假设:

$$H_0: P_i = \frac{1}{6}, i = 1, 2, \dots, 6, \quad H_1: \exists i, \text{使} P_i \neq \frac{1}{6}.$$

解法 1: 根据公式(1.6)计算检验的概率值, 记抽样频率与目标概率的平均距离为

$$A = \frac{1}{6} \left\{ \left| \frac{12}{100} - \frac{1}{6} \right| + \left| \frac{25}{100} - \frac{1}{6} \right| + \left| \frac{13}{100} - \frac{1}{6} \right| + \left| \frac{12}{100} - \frac{1}{6} \right| + \left| \frac{23}{100} - \frac{1}{6} \right| + \left| \frac{15}{100} - \frac{1}{6} \right| \right\}$$

$$\text{检验概率值为: } p = P \left\{ (n_1, n_2, \dots, n_6) | (n_1, n_2, \dots, n_6) \in \Omega, \frac{1}{6} \sum_{i=1}^6 \left| \frac{n_i}{100} - \frac{1}{6} \right| \geq A \right\}$$

用 Python 语言, 得概率值 $p = 0.040$ 。由于 $p < 0.05$, 拒绝原假设。

3. 几种检验方法的效率比较

带判别函数的假设检验方法是一种一般性的方法, 不同的判别函数构成不同的检验方法。上一小节的方法只是一个特例, 传统的二项分布检验方法、拟合优度方法、和似然比检验方法也是这个方法的特例。本文还列举了其它判别函数的检验方法。

二项分布检验有一种精确分布计算方法。设一总体为 0-1 分布, 其中成功的概率为 p 。对总体进行 m 次重复抽样, 其中成功 m_1 次, 不成功 $m - m_1$ 次, 可以得到 m_1 服从二项分布 $B(m, p)$ 。

现进行假设检验, $H_0: p = p_0, H_1: p \neq p_0$ 。给定显著水平 α , 拒绝域为 $B(m, p)$ 分布中两端概率各为 0.5α 的区间。因此检验的概率值为 $B(m, p)$ 分布中 m_1 两端中概率小的一端的两倍数, 值,

$$p = 2 \min \{ P\{0 \leq k \leq m_1\}, P\{m_1 \leq k \leq m\} \} \quad (3.1)$$

当概率值小于显著水平时, 就拒绝原假设, 否则就接受(不拒绝)原假设。

这样定义的拒绝域有两个缺点。首先, 由于二项分布左右不对称, 这种两端拒绝域的概率各为 0.5α 的设计不太合理, 因为它的接受域的长度不是最短的。其次, 这种定义不能推广的多个状态分布的假设检验, 因为多项分布是多维的, 不能定义各方向的极端区域概率相等。

本文对这种检验方法做了修改, 并推广到 K 个状态分布检验(1.2)的情形。二项分布检验只是一个特例, 在此不做专门描述。先待定一个正常数 c , 记

$$S(c) = \left\{ (n_1, n_2, \dots, n_K) \mid (n_1, n_2, \dots, n_K) \in \Omega, \frac{m!}{\prod_{i=1}^K n_i!} \prod_{i=1}^K p_{i0}^{n_i} \leq c \right\} \quad (3.2)$$

$S(c)$ 是在样本空间中, 概率小于等于 c 的点组成的集合。给定显著水平 α , 假设检验的拒绝域定义为

$$\text{Rej}(\alpha) = \bigcup_{c>0} \{S(c) \mid P\{S(c)\} < \alpha\} \quad (3.3)$$

拒绝域是在 $\text{Rej}(c)$ 中概率小于 α 的最大集合, 而接受域就是 $\Omega \setminus \text{Rej}(c)$ 。

上述方法的判别函数为样本的概率

$$P(N, P_0) \triangleq \frac{m!}{\prod_{i=1}^K n_i!} \prod_{i=1}^K p_{i0}^{n_i} \quad (3.4)$$

检验的概率值为概率小于等于抽到的样本概率的所有样本构成的集合的概率。

$$p = P \left\{ (n_1, n_2, \dots, n_K) \mid (n_1, n_2, \dots, n_K) \in \Omega, \frac{m!}{\prod_{i=1}^K n_i!} \prod_{i=1}^K p_{i0}^{n_i} \leq \frac{m!}{\prod_{i=1}^K m_i!} \prod_{i=1}^K p_{i0}^{m_i} \right\} \quad (3.5)$$

当概率值 p 小于显著水平时, 抽到的样本落在拒绝域里, 因此拒绝原假设; 否则就接受原假设。

例 1 的解法 2: 用公式(2.5)计算检验的概率值。记抽样样本的概率为

$$A = \frac{100!}{12!25!13!12!23!115!} \left(\frac{1}{6}\right)^{100}$$

$$\text{检验的概率值为: } p = P \left\{ (n_1, n_2, \dots, n_6) \mid (n_1, n_2, \dots, n_6) \in \Omega, \frac{100!}{\prod_{i=1}^6 n_i!} \left(\frac{1}{6}\right)^{100} \leq A \right\}$$

用 Python 语言, 得概率值 $p = 0.087$ 。由于 $p \geq 0.05$, 接受原假设。

下面说明似然比检验方法是带判别函数检验的一个特例。似然比定义为无条件的最大似然值除以在零假设限制下的最大似然值。在 K 个状态分布检验(1.2)中, 给定样本点 N ,

$$\sup \left\{ \prod_{i=1}^K p_i^{n_i} \right\} = \prod_{i=1}^K \left(\frac{n_i}{m}\right)^{n_i}, \quad \sup_{P_0} \left\{ \prod_{i=1}^K P_i^{n_i} \right\} = \prod_{i=1}^K (p_{i0})^{n_i} \quad (3.6)$$

$$S(c) \triangleq \left\{ (n_1, n_2, \dots, n_K) \mid (n_1, n_2, \dots, n_K) \in \Omega, \prod_{i=1}^K \left(\frac{n_i}{m}\right)^{n_i} / \prod_{i=1}^K p_{i0}^{n_i} \geq c \right\} \quad (3.7)$$

$S(c)$ 是在样本空间中, 似然比大于等于 c 的点组成的集合。检验的思路是, 似然比越大, 与无条件最大似然值相比, 在零假设条件下最大似然值减少越多, 零假设越不可能是真实的。

给定显著水平 α , 假设检验的拒绝域定义为

$$\text{Rej}(\alpha) = \bigcup_{c>0} \{S(c) \mid P\{S(c)\} < \alpha\} \quad (3.8)$$

拒绝域就是在 $\text{Rej}(c)$ 中, 概率小于 α 的最大集合, 而接受域就是 $\Omega \setminus \text{Rej}(c)$ 。

在似然比检验中, 判别函数为似然比,

$$P(M, P_0) \triangleq \prod_{i=1}^K \left(\frac{m_i}{m} \right)^{m_i} / \prod_{i=1}^K p_{i0}^{m_i} \quad (3.9)$$

检验的概率值为似然比大于等于抽到的样本的似然比的所有样本构成集合的概率。

$$p = P \left\{ (n_1, n_2, \dots, n_K) \mid (n_1, n_2, \dots, n_K) \in \Omega, \frac{\prod_{i=1}^K \left(\frac{n_i}{m} \right)^{n_i}}{\prod_{i=1}^K p_{i0}^{n_i}} \geq \frac{\prod_{i=1}^K \left(\frac{m_i}{m} \right)^{m_i}}{\prod_{i=1}^K p_{i0}^{m_i}} \right\} \quad (3.10)$$

当概率值 p 小于显著水平时, 抽到的样本点落在拒绝域里, 拒绝原假设, 否则就接受原假设。

例 1 的解法 3: 用公式(2.10)计算检验的概率值。记抽样样本的似然比为

$$A = \left(\frac{12}{100} \right)^{12} \left(\frac{25}{100} \right)^{25} \left(\frac{13}{100} \right)^{13} \left(\frac{12}{100} \right)^{12} \left(\frac{23}{100} \right)^{23} \left(\frac{15}{100} \right)^{15} / \left(\frac{1}{6} \right)^{100}$$

$$\text{检验的概率值: } p = P \left\{ (n_1, n_2, \dots, n_6) \mid (n_1, n_2, \dots, n_6) \in \Omega, \prod_{i=1}^6 \left(\frac{n_i}{100} \right)^{n_i} / \left(\frac{1}{6} \right)^{100} \geq A \right\}$$

用 Python 语言, 得概率值 $p = 0.087$ 。由于 $p \geq 0.05$, 接受原假设。

下面说明 X^2 拟合优度检验是带判别函数假设检验的一个特例。在 K 个状态分布检验(1.2)中, 拟合优度检验的 Pearson X^2 统计量为

$$X^2 = \sum_{i=1}^K \frac{(n_i - mp_{i0})^2}{mp_{i0}} = m \sum_{i=1}^K p_{i0} \left(\frac{n_i}{m} / p_{i0} - 1 \right)^2 \quad (3.11)$$

统计量可转化为该检验的判别函数

$$rd^2(N, P_0) = \sum_{i=1}^K p_{i0} \left(\frac{n_i}{m} / p_{i0} - 1 \right)^2 \quad (3.12)$$

判别函数是样本频率与目标概率相对差的二阶矩。样本的判别函数越大, 样本频率与目标概率相差就越大, 样本与零假设就越相悖。记,

$$S(c) \triangleq \left\{ (n_1, n_2, \dots, n_K) \mid (n_1, n_2, \dots, n_K) \in \Omega, \sum_{i=1}^K p_{i0} \left(\frac{n_i}{m} / p_{i0} - 1 \right)^2 \geq c \right\} \quad (3.13)$$

给定显著水平 α , 假设检验的拒绝域定义为

$$\text{Rej}(\alpha) = \bigcup_{c>0} \{S(c) \mid P\{S(c)\} < \alpha\} \quad (3.14)$$

拒绝域就是在 $\text{Rej}(c)$ 中, 概率小于 α 的最大集合, 而接受域就是 $\Omega \setminus \text{Rej}(c)$ 。

检验的概率值为判别函数大于等于抽到的样本的判别函数的所有样本构成集合的概率,

$$p = P \left\{ (n_1, n_2, \dots, n_K) \mid (n_1, n_2, \dots, n_K) \in \Omega, \sum_{i=1}^K p_{i0} \left(\frac{n_i}{m} / p_{i0} - 1 \right)^2 \geq \sum_{i=1}^K p_{i0} \left(\frac{m_i}{m} / p_{i0} - 1 \right)^2 \right\} \quad (3.15)$$

当概率值 p 小于显著水平时, 抽到的样本点落在拒绝域里, 拒绝原假设, 否则就接受原假设。

例 1 的解法 4: 用公式(2.15)计算检验的概率值。记样本频率与目标概率相对差的二阶矩为

$$A = \frac{1}{6} \left\{ \left(\frac{12}{100} / \frac{1}{6} - 1 \right)^2 + \left(\frac{25}{100} / \frac{1}{6} - 1 \right)^2 + \left(\frac{13}{100} / \frac{1}{6} - 1 \right)^2 + \left(\frac{12}{100} / \frac{1}{6} - 1 \right)^2 + \left(\frac{23}{100} / \frac{1}{6} - 1 \right)^2 + \left(\frac{15}{100} / \frac{1}{6} - 1 \right)^2 \right\}$$

$$\text{检验的概率值: } p = P \left\{ (n_1, n_2, \dots, n_6) \mid (n_1, n_2, \dots, n_6) \in \Omega, \frac{1}{6} \sum_{i=1}^6 \left(\frac{n_i}{100} / \frac{1}{6} - 1 \right)^2 \geq A \right\}$$

用 Python 语言, 得概率值 $p = 0.071$ 。由于 $p \geq 0.05$, 接受原假设。

带判别函数检验方法的函数形式是很多的。除了上述四种判别函数, 本文补充两个, 都是反映样本的分布频率与目标概率的差距。例如, 对假设检验(1.2), 判别函数可定义为样本的分布频率与目标概率的差距的二阶矩

$$d^2(N, P_0) \triangleq \sum_{i=1}^K p_{i0} \left(\frac{n_i}{m} - p_{i0} \right)^2 \quad (3.16)$$

也可以定义为样本的分布频率与目标概率的差距的最大值

$$d^M(N, P_0) \triangleq \max_{1 \leq i \leq K} \left\{ \left| \frac{n_i}{m} - p_{i0} \right| \right\} \quad (3.17)$$

不同的检验方法存在检验效率的差异, 其中一个检验效率的指标是概率值。面对同一个样本, 如果方法 A 比方法 B 有更小的概率值, 方法 A 的检验效率就比方法 B 更高。这是因为较小的概率值更容易小于显著水平, 也就更容易拒绝原假设。

针对例 1 这个案例, 我们看到, 方法 1 的概率值最小(为 0.040), 概率值小于 0.05, 检验拒绝了原假设。其次是方法 4、方法 3 和方法 2, 其概率值分别为 0.071、0.087 和 0.087。这三个检验的概率值都大于 0.05, 不能拒绝原假设。可以看到, 以样本的分布频率与目标概率的距离的平均值为判别函数的检验方法, 比传统的三个检验方法都更有检验效率。

4. 结论

本文引进了一种带判别函数的单样本有限状态分布检验, 并给出了基于精确分布的检验计算方法。在本文方法中, 给出不同的判别函数, 检验可表示不同的具体检验方法, 其中几个经典的检验方法, 例如 X^2 拟合优度检验、二项检验和似然比检验都是本文检验方法在不同判别函数下的特例。

主流的分布检验理论都是借用极限分布来处理的。在这个方法中, 真实分布与极限分布的差异会导致检验结果出现误差, 且这种误差并不能在检验中被衡量得到。由于计算机和软件的快速发展, 使用精确分布处理假设检验变得越来越可行。

精确分布方法可带来两大改进。首先, 该方法可以避免由分布差异导致的检验误差问题, 特别是在样本量不是很大的时候。其次, 在借用极限分布处理分布检验时, 不同的判别函数会导致不同的极限分布, 甚至有可能连极限分布都不存在。在这种情况下, 判别函数的选取会受制于它的极限分布的存在性和表达形式。而在使用精确分布处理假设检验时, 精确分布就是总体分布, 与判别函数的选取无关。在这种情况下, 我们就可以选择最大程度地表现样本频率和目标概率差异的判别函数, 以提高假设检验的效率。

本文并没有从理论上给出上述六个检验方法的优劣, 因此未能得出一般性结论。但针对例 1 这个案例, 本文分析显示, 以样本的分布频率与目标概率的距离的平均值为判别函数的检验方法, 比传统的检验方法更有效率, 这类方法值得我们进一步研究。

参考文献

- [1] Fisher, R.A. (1928) On a Property Connecting the χ^2 Measure of Discrepancy with the Method of Maximum Likelihood. *Atti de Congresso Internazionaledei Matematici*, **6**, 94-100.
- [2] Watson, G.S. (1958) On Chi-Square Goodness-of-Fit Tests for Continuous Distributions. *Journal of the Royal Statistical Society, Series B (Methodological)*, **20**, 44-72. <https://doi.org/10.1111/j.2517-6161.1958.tb00274.x>
- [3] 陈希孺. 高等数理统计学[M]. 合肥: 中国科学技术大学出版社, 1999: 230-242.
- [4] Brenner, D.J. and Quan, H. (1990) Exact Confidence Limits for Binomial Proportions—Pearson and Hartley Revisited. *Journal of the Royal Statistical Society. Series D (The Statistician)*, **39**, 391-397. <https://doi.org/10.2307/2349083>
- [5] Clopper, C.J. and Pearson, E.S. (1934) The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. *Biometrika*, **26**, 404-413. <https://doi.org/10.1093/biomet/26.4.404>
- [6] Krishnamoorthya, K., Thomsonb, J. and Caia, Y. (2004) An Exact Method of Testing Equality of Several Binomial Proportions to a Specified Standard. *Computational Statistics & Data Analysis*, **45**, 697-707. [https://doi.org/10.1016/S0167-9473\(03\)00095-1](https://doi.org/10.1016/S0167-9473(03)00095-1)
- [7] Neyman, J. and Pearson, E.S. (1928) On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference—Part II. *Biometrika*, **20A**, 263-294. <https://doi.org/10.1093/biomet/20A.3-4.263>
- [8] 杨振海. 拟合优度检验[M]. 合肥: 安徽教育出版社, 1994: 50-51.