第 42 卷 第 6A 期
2015 年 6 月

计 算 机 科 学
Computer Science

Vol. 42 No. 6A
June 2015

多重假设检验及其在大数据特征降维中的应用

潘 舒 祁云嵩

(江苏科技大学计算机科学与工程学院 镇江 212003)

摘 要 现有的特征降维方法大致可分为特征提取和特征选择。在特征提取过程中,数据中的原始特征通过某些数据变换被映射到一个低维空间。提取出的特征尽管与原始特征相关,但不再具有原始特征的物理意义,即特征提取改变了原始数据的表达形式。与特征提取不同,特征选择则在原有的特征集中选择一个子集,选择出的特征子集中不再含有与数据分析任务相关性不大或冗余的那部分特征,其结果可能引起信息丢失。因而现有的数据降维方法几乎都不是保真降维,其降维后的数据仅适合特定的后续数据分析任务,因而只能算是特定数据分析任务的前期数据预处理。从多重假设检验方法的角度分析了高维数据保真降维的方法及研究的关键所在。

关键词 特征选择,降维,多重假设检验

中图法分类号 TP18 文献标识码 A

Multiple Hypothesis Testing and its Application in Feature Dimension Reduction

PAN Shu QI Yun-song

(School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

Abstract The existing feature dimension reduction methods can roughly be categorized into two classes: feature extraction and feature selection. In feature extraction problems, the original features in the measurement space are initially transformed into a new dimension-reduced space via some specified transformation. Although the significant variables determined in the new space are related to the original variables, the physical interpretation in terms of the original variables may be lost. So, feature extraction will change the description of the original data. Unlike feature extraction, feature selection aims to seek optimal or suboptimal subsets of the original features by preserving the main information carried by the complete data to facilitate future analysis for high dimensional problems. Often, the selected features are a subset of the original features, and those insignificant and redundant features may be discarded. It is worth mentioning that almost all of the existing dimensionality reduction methods are not high fidelity methods. The result of these methods is only suitable for specific subsequent data analysis tasks, which is only a particular task under the preprocess. In this paper, with the technique of multiple hypothesis testing, we studied the dimensionality high fidelity reduction problem. The processing results can save all the useful information and eliminate the irrelevant features from the original data.

Keywords Feature selection, Dimension reduction, Multiple hypothesis testing

1 前言

在当今信息和知识经济时代,人类研究的科学和社会问题更加高深复杂,更加庞大,有效地收集和分析数据以提取信息和知识变得更加重要和紧迫。发达的信息技术和高性能计算机的使用使得收集、储存、传输数据及科学计算更加便捷,所搜集数据中维数也呈几何级数的速度增长,其数据维数经常远高于样本量的个数。虽然高维数据中蕴藏着丰富的信息,但其数据的处理却面临着维数灾难(Curses of Dimensionality)问题。Bellman 等人曾研究了高维空间上的函数优化、函数逼近及数值积分等,其研究结果表明:要达到给定的精度,运算复杂性的阶数是数据维数 D 的指数函数[1];在统计估计过程中,为达到相同的估计精度,所需的样本数随维数的增加而呈指数增长。另外,许多经典的低维数据处理方法,如回归分析、主成分分析、独立成分分析等,在处理高维数据时存在本质的困难[2-4]。例如,膨胀的维数导致计算量迅速上升,高维特性导致样本数量相对较少,从而使得某些统计上的渐近性质受到破坏,传统方法在处理高维数据时不满足稳健性要求等。

降维是用来克服"维数灾难"和高维数据模型化的一种数据处理技术,是用来解决这一问题的有效手段之一。它可通过对离散数据集合的分析来探求嵌入在高维数据空间中本征低维流形的不同样式,寻求事物的本质规律性。特征降维(Feature Dimension Reduction)是一个从初始高维特征集合

本文受国家自然科学基金(61471182),江苏省高校自然科学基金(13KlB520003)资助。

潘 舒(1981—),女,硕士,讲师,主要研究方向为智能信息处理,E-mail:may-nap@sina.com;祁云嵩(1967—),男,博士,教授,主要研究方向为智能信息处理、模式识别等。

中选出低维特征集合，根据一定的评估准则最优化缩小特征空间的过程，通常作为机器学习等数据分析处理的预处理步骤。特征降维自20世纪70年代以来就得到了广泛的研究。大量研究实践证明，特征降维能有效地消除无关和冗余特征，提高挖掘任务的效率，改善预测精确性等学习性能[5]。然而，数据在数量和维度上的剧增趋势也对特征降维算法提出了更加严峻的挑战。现有数据降维方法大多以获取原始数据在特定任务下的最佳特征子集为目的，其结果是抛弃那些“作用不大”（并非没有作用）或冗余的特征。这并不能保证原始数据降维处理后不失真，因而其降维结果只能适用于特定的原始任务。本课题着力于获取高维数据中全部（期望）的“有用”特征，去除其“无关”特征，从而使得降维后的数据最大限度地不失真。

2 特征降维方法

通常，高维数据存在以下几方面的问题：大量的特征；许多与给定任务无关或与类别仅有微弱相关度的特征；许多对于给定任务冗余的特征，如特征之间存在强烈的相关度；噪声数据。特征降维是一种通过降低特征维度从而提高给定任务效率的方法，可以分为特征选择和特征提取两种降维方式。特征选择是通过一定的算法在全部特征中选择出期望的最佳特征子集，而特征提取则是通过原始特征的变换、组合等手段在原有的特征基础上构建新的特征子集，如传统的Fisher判断[6]、投影寻踪（Projection Pursuit）[7,8]等。尽管特征提取所形成的新的特征子集具有较好的鉴别能力，但这些新的特征由于是由原始特征线性或非线性组合而成，因此失去了其原有的物理意义。在某些场合，原始特征的物理意义显得更为重要。例如，在基因微阵列数据分析中，如果研究人员更关心的是个体基因是否在不同类别的样本中差异表达，则应该选用特征选择进行高维数据的特征降维处理。此外，从计算性能上考虑，特征抽取需要更复杂的计算。所以，在高维数据分析过程中，往往更需要特征选择或先进行特征选择后再进行特征提取。综上，在特征降维中，特征选择显得更为重要。

特征选择方法可分为过滤法（Filter Methods）[9]、缠绕法（Wrapper Methods）[10]以及嵌入法（Embedded Methods）[11]。过滤法根据特征的固有特性在原始数据的基础上不经过复杂的学习算法直接确定所选择的特征子集，缠绕法则根据一定的学习算法通过评估特征子集在学习算法中的性能来决定最终所选择的特征子集；嵌入法则是在方法模型建立过程中即开始对特征进行评估选择。相比较而言，缠绕法和嵌入法需要更复杂的计算，在高维数据处理时往往要先经过过滤法对特征进行预选择，并且这两种方法最终所得到的特征子集往往是全部相关特征中最适合所用算法的一个子集。因此，从数据降维的角度出发，基于过滤法的特征选择方法更实用、更有效。

现有代表性的基于过滤法的特征选择方法有Exhaustive Search，Branch-and-Bound Search、Best Individual Features、Sequential Forward Selection（SFS）、Sequential Backward Selection（SBS）、Sequential Forward Floating Search（SFFS）、Sequential Backward Floating Search（SBFS）等。此外，基于粗糙集的特征选择方法近年来得到了国内外众多研究学者的关注[12-14]。该方法的一个核心思想就是在保持数据集中信息量、近似分布等度量不变的情况下，删除数据集中的冗余特征，以达到降维的目的。用于粗糙集特征选择的典型方法有分辨矩阵和启发式算法。分辨矩阵是一种穷举式搜索方法，已被证明是一个NP-Hard问题，所以在很多实际应用问题中往往采用启发式算法。为了进一步提高启发式算法的效率，亦有很多学者提出了诸如粗糙集加速器[13,14]、增量式启发函数[15,16]等。基于过滤法的特征选择方法的研究重点是使用不同搜索策略选择最佳的特征子集。从模式识别的角度出发，特征选择是选择尽可能少的特征以求得最大的分类准确率。在这些特征选择过程中，往往抛弃那些鉴别能力不强的或冗余的特征。如果换一种特征选择方法，其选择结果不尽相同。然而，那些被抛弃的特征并非“无用”的特征。从数据降维的角度出发，为了保证降维后的数据不丢失原有的信息，必须研究相关的能保留全部（期望的）有用特征的特征选择方法。

3 假设检验问题

特征选择问题可以描述为假设检验问题，即特定的数据特征是否为数据所描述问题相关的特征。假设检验问题是统计推断和决策的基本形式之一，其核心内容是利用样本所提供的信息对关于总体的某个假设进行检验。对于这一问题，经典统计学派和贝叶斯学派有不同的处理方法和检验法则。经典统计学派的假设检验主要是运用概率反证法进行推断，它主要有两种方法：一种是Gossett于1908年提出的p值检验，一种是Nehman和Pearson分别于1928年和1933年提出的固定水平检验。p值检验的基本思想是：选择一个检验统计量，在假定原假设为真时计算此检验统计量的值及对应的概率p，若该p值小于事先给定的显著水平α，则拒绝原假设H_0，若该p值大于事先给定的显著性水平α，则不拒绝原假设H_0。固定水平检验的基本思想是：选择一个检验统计量，在事先给定的显著性水平α下确定拒绝域，当检验统计量的值落入拒绝域时，则拒绝原假设H_0，否则接受原假设H_0。相对于经典统计学派的假设检验方法，贝叶斯学派的检验方法是直截了当的。它是在获得后验分布后，直接计算原假设H_0和对立假设H_1的后验概率P_0和P_1，并通过比较两个后验概率的大小决定假设检验的结果：如果$P_0>P_1$，则接受原假设H_0；否则接受对立假设H_1（拒绝原假设）。相关研究表明，经典学派的p值与贝叶斯学派的后验概率大不相同。在正态分布的前提下，当经典方法得到的p值在$0.001\sim0.1$之间时，贝叶斯方法得到的原假设H_0的后验概率却很大，即此时经典方法倾向于拒绝原假设，而贝叶斯方法则倾向于接受原假设[17]。在高维数据中，正常情况下大多数特征为无关特征，假设检验中的原假设一般是“所选择的特征为无关特征”。为了保证降维后的数据不失真，在检验过程中应该尽可能拒绝原假设。所以，在特征降维时，更应该使用经典的假设检验方法。

对于经典的假设检验，其方法是分析样本数据并通过统计方法计算其p值（与假设相关的概率）。如果该p值小于某一显著性阈值，则可拒绝相应的原假设。然而，当众多的假设一起检验时，总体错误率将随着原假设数量的增多而急剧上升。所以，多重假设检验流程（Multiple Testing Procedure，MTP）[18]被用来控制总体错误率。早期的多重假设检验控制

的是假设检验流程中(至少)出现一个(第Ⅰ类)错误的概率,即控制的是簇错误率(Family-Wise Error Rate,FWER)[19]。为说明其意义,表1列出了多重假设检验的4种结果。

表1 多重假设检验4种结果

真实情况	不拒绝原假设H_0数	拒绝原假设H_0数	总数
原假设H_0为真	U	V	m_0
原假设H_0为假	T	S	m_1
总数	W=m-R	R	m

FWER定义为拒绝真实无效假设的个数大于或等于1的概率(记作$P(V\geqslant 1)$),通常使用Bonferroni法对其进行检验,即对每一个假设都在显著性水平α/m下进行检验,保证$FWER=P(V\geqslant 1)$小于或等于事先给定的检验水准σ。然而,这种控制方式在原假设数量较大的情况下的控制结果极为保守[20],其FWER的实际意义也不够直观和容易理解。1995年,Benjamini和Hochberg提出用错误发现比例的期望值(False Discovery Rate,FDR)[20]指标代替FWER。该方法主要源于这样的一个事实:当原假设的数量比较大时,人们可以容忍更多的错误发现数量。例如,在10个假设中出现5个错误当然太多,但人们也许可以接受100个假设中出现5个错误的事实。在高维数据中特征数量众多,因而FDR控制指标的使用能大大增强假设检验的功效。

根据表1,FDR定义如下:

$$FDR=\begin{cases}E(V/R), & R\neq 0\\ 0, & R=0\end{cases}$$

其中,$E(\cdot)$为数学期望。FDR的含义是阳性检验结果中判断错误的比例。FDR作为假设检验错误率的控制指标,其控制值可以根据需要灵活选取;而FWER作为控制指标的假设检验,取值则较为固定,通常定为0.05。此外,FDR的意义也较明确,可以作为筛选出的差异特征变量的评价指标,FWER则主要是用来控制Ⅰ类错误。当所有被拒绝的假设为真($V\geqslant 1$)时,控制FDR和控制FWER等价;当$m_0<m$时,控制FDR相当于FWER的弱控制。

3.1 FDR控制

FDR控制是指决定一个显著性水平的界值,从而使检验结果的FDR被限制在某一固定水平[21-25]。Benjamini和Hochberg首次提出的FDR控制流程(BH流程方法)采用逐步向上(Step-Up)的控制方法[20],分两步进行:首先将所有检验的p值进行排序,即$p_{(1)}\leqslant p_{(2)}\leqslant\cdots p_{(m)}$;然后逐步后退比较$p_{(i)}\leqslant i\alpha/m(i=m,m-1,m-2,\cdots,1)$,取第一个满足条件的$p_{(k)}(k\geqslant 1)$为拒绝原假设的阈值,即拒绝原假设$H_{(j)}(j=1,2,\cdots,k)$。理论上可以证明,在该情况下可以将FDR控制在$\alpha(0\leqslant\alpha\leqslant 1)$水平之下。该方法需要满足各假设检验的特征变量间相互独立的条件。在此基础上,Yekutieli和Benjamini于1999年给出了一种改进的方法[26],其思想是利用重复抽样的方法来计算p值,可以在特定的相关条件下控制FDR值,但其估计的FDR值略为保守。同年,Benjamini和Liu提出了一种逐步向下(Step-Down)的控制方法[27],过程与BH基础方法相近,只是对$p_{(k)}$的控制方法不同。2000年Benjamini和Hochberg提出了两阶段的FDR控制以改进原有方法的保守性[28];2001年Benjamini和Yekutieli对算法进行了进一步的改进,改进后的算法可用于不同检验变量间独立和相关条件下的FDR的控制[29]。Benjamini和Hochberg在2006年提出了一种自适应线性向上的控制方法(Adaptive Llinear Step-Up,ALSU)[30],这种方法的特点是在不同的显著水准下两次使用上述介绍的基础过程,然而,在变量相关条件下这一方法的检验效能仍较为保守。

要将FDR控制在一个期望水平以下,几乎所有的方法都在试图精确地估计FDR[31]。FDR估计是指在某一"有统计学意义"阈值情况下所有被拒绝假设的错误发现率估计。对此,可以从贝叶斯角度给出解释,近期对于FDR的研究也多数在此框架下进行。贝叶斯定义使用两分量模型来构建p值的分布函数:$F(p)=\pi_0F_0(p)+\pi_1F_1(p)$,其中$\pi_0$为真实无效假设所占总检验次数的比例,$F_0$为无效假设下$p$的分布函数,$\pi_1$为真实有效假设所占总检验次数的比例,$F_1$为有效假设下$p$的分布函数。根据贝叶斯公式,可得

$$FDR=\Pr\{\text{null}\mid p\}=\frac{\pi_0F_0(p)}{F(p)}$$

若用概率密度函数代替分布函数,则可以得到局部FDR(记作Lfdr):

$$f(p)=\pi_0f_0+\pi_1f_1(p)$$

$$Lfdr=\Pr\{\text{null}\mid p\}=\frac{\pi_0f_0}{f(p)}$$

其意义为在某p值时,错误拒绝该假设的概率。Lfdr的提出,使得多重假设检验能够估计出任意一次检验出错的概率,通常情况下有$FDR\leqslant Lfdr$。

现有的FDR估计方法中,应用较多的还有q-value法、SAM法以及经验贝叶斯法等3种。q-value方法于2003年由Storey提出[32],其基本思想是在最初的FDR控制过程的基础上估计成立的原假设占全部假设总数的比例π_0,以提高原有方法的检验效能。SAM方法是Tusher等人2001年提出的一种分析微阵列数据中基因差异表达显著性的方法(Significance Analysis of Microarray)[33]。该方法采用Permutation方法估计错误发现率FDR。Efron注意到在FDR的估计方法中,许多方法假定无效假设的p值服从均匀分布,这在"理想情况下"是成立的。然而,由于实际数据的复杂性且通常不能严格满足所要求的条件(如t检验要求的方差齐性、正态性),导致在无效假设下p值不再服从均匀分布,使FDR的估计出现偏差,即使利用Permutation方法也并非总能得到无效假设下的真实分布[34],由此得到的FDR也不一定准确。许多估计方法,如q-value[32]、kerfdr[35]、locfdr[36]等方法皆依赖于该均匀分布的假设,SAM则是基于Permutation抽样的FDR估计方法。Efron提出采用经验贝叶斯法直接对无效假设下检验统计量的分布进行估计,如对z统计量的均数和方差进行经验估计,而不用标准正态分布进行检验,重新计算p值[34]。

对于FDR的控制与估计,使用者主要关心其估计的无偏性、稳定性、检验效能和不同数据结构(如相关性)对FDR控制或估计的影响[37-41]。仿真实验证实,在大部分理想情况下FDR的各类控制方法能够将错误控制在指定水平下(BH控制方法偏于保守),并且是无偏估计。在特征变量间存在"弱相关"条件下,SAM、q-value等大部分方法对FDR的估计依然是稳健的;在变量存在简单正相关条件下,BH基础算法依然保持变量独立条件下的性质。但是在"任意相关"条件下,仿真实验证实SAM和q-value等方法对π_0和FDR的估计将产生较大的变异和偏差,而此时ALSU方法则显现出较好的

稳定性[42]。经验贝叶斯法在理论上能够对无效假设进行更好的拟合,解决任意相关对 FDR 估计的影响,但实际数据分析发现表明,这种方法有时并不能得到理想的结果[34]。如在对一项乳腺癌与正常样本的微阵列对照数据的分析中,将 FDR 控制在一定水平,使用 locfdr 算法未发现有差异基因,而使用简单的 BH 算法却发现了 107 个"差异表达基因"[34]。需要注意的是,许多方法如 q-value、locfdr、kerfdr 等用于实际复杂数据分析时,可能由于无法满足适用条件使 π_0 和 FDR 的估计明显超出合理范围[34]。

3.2 需要解决的关键问题

尽管多重假设检验方法流程在理论上已相对成熟,但由于数据样本本身的特殊性,其实用性还存在诸多瓶颈问题。目前,相关研究热点主要集中在以下几个方面。

1)小样本数据集中 p 值估算问题

这里的小样本特指样本绝对数量较小。在多重假设检验方法中,使用得最多的是 t 检验。当样本数量较少时,t 检验的方差统计极不稳定,从而使特征变量 p 值计算的准确性难以符合要求[33]。Permutation 方法被认为是解决这一问题较好的选择[32]。设数据集中某个特征变量的 t 检验统计量为 t_i,对全部样本通过全排列的方式进行重新组合后 M 种可能情况下的统计量分别为 $\{t_i^1, t_i^2, \ldots, t_i^M\}$,设这些统计量中大于 t_i 的个数量为 N,则相应的特征的 p 值计算为 N/M。当样本总数很少时,其 M 值也相对很小,从而导致较小的 p 值,计算不稳定。所以,在样本绝对数量较少的情况下,相关特征变量的 p 值计算仍然是一个值得研究的问题。

2)零假设比例估算问题

多重假设检验过程是基于 p 值计算的,因此 p 值的准确性最终影响 FDR 估计的准确性。由于 p 值的计算依赖于一定的假设,无论使用什么样的数据模型,其实际成立与不成立的原假设比例估算的准确性直接影响 p 值计算的准确性,进而影响 FDR 控制方法的功效。文献[43]的研究表明,如果知道原假设的真实比例,还可以进一步提高假设检验的功效。大多数文献研究在估算原假设比例时均假设非显著性差异特征变量的 p 值服从均一分布。通过大量的比较研究,文献[44]指出这一假设常常会使得零假设的比例估算偏高。此外,原假设比例的估算也直接影响 FDR 的估算[34],然而,这一问题还没有得到很好的解决。

3)变量相关问题

高维数据集中特征变量间的相关性严重影响多重假设流检验的结果。文献[42]的仿真研究表明:Benjamin 等人的 FDR 流程控制[20,28,29,45]、Storey 的 q-value 方法[32]以及 Tusher 等人的 SAM 方法[33]均不能在特征变量相关的情况下充分地控制假设检验的 FDR;尽管 Benjamini 和 Hochberg 的 adaptive 方法[30]能适应相关特征变量间的假设检验,但其 FDR 控制结果仍然较为保守。因此,如何提高特征变量相关情形下的多重假设检验效果仍是众多学者的研究焦点。

4)无监督检验问题

现有的假设检验问题基本是对两类或多类样本中特征变量差异的显著性进行检验。然而,在科学研究或实际应用中,经常需要对无类别标识的数据进行分析处理。如何将多重假设检验用于无类别标识的数据分析,即无监督多重假设检验问题的研究在特征降维处理中有一定的实际意义。

结束语 现所有的数据降维方法几乎都不是保真降维,其降维后的数据仅适合特定的后续数据分析任务,因而只能算是特定数据分析任务的前期数据预处理。本文以多重假设检验方法为出发点,旨在探索一类高保真数据降维方法,其降维结果致力于保留原始数据中的全部(期望的)原始特征,最大限度地剔除无关特征。该方法的研究对大数据清洗、存储等有实际意义。

参 考 文 献

[1] Bellman, Richard. Adaptive Control Processes: A Guided Tour [M]. Princeton University Press, 2000

[2] 于玲,吴铁军. LS-Ensem:一种用于回归的集成算法[J]. 计算机学报,2006,29(5):719-726

[3] 钱叶魁,陈鸣,叶立新,等. 基于多尺度主成分分析的全网络异常检测方法[J]. 软件学报,2012,23(2):361-377

[4] 黄雅平,罗四维,陈恩义. 基于独立分量分析的虹膜识别方法[J]. 计算机研究与发展,2003,40(10):1451-1457

[5] Jie H. Survey on feature dimension reduction for high-dimensional data[J]. Application Research of computers, 2008, 9(8)

[6] 杨静,于旭,谢志强. 改进向量投影的支持向量预选取方法[J]. 计算机学报,2012,35(5):1002-1010

[7] 宋枫溪,高秀梅,刘树海,等. 统计模式识别中的维数削减与低损降维[J]. 计算机学报,2005,28(11):1915-1922

[8] Huber P. Projection pursuit[J]. The annals of Statistics, 1985, 13(2):435-475

[9] 徐峻岭,周毓明,陈林,等. 基于互信息的无监督特征选择[J]. 计算机研究与发展,2012,49(2):372-382

[10] Wang H, Das S R, Suh J W, et al. A learning-based wrapper method to correct systematic errors in automatic image segmentation: Consistently improved performance in hippocampus, cortex and brain segmentation[J]. NeuroImage, 2011, 55(3): 968-985

[11] Cheng M, Fang B, Pun C M, et al. Kernel-view based discriminant approach for embedded feature extraction in high-dimensional space[J]. Neurocomputing, 2011, 74(9): 1478-1484

[12] Qian Y, Zhang H, Sang Y, et al. Multigranulation decision-theoretic rough sets[J]. International Journal of Approximate Reasoning, 2014, 55(1): 225-237

[13] Qian Yuh-ua, Liang Ji-ye, Pedrycz W, et al. Positive approximation: an accelerator for attribute reduction in rough set theory [J]. Artificial Intelligence, 2010, 174: 597-618

[14] 王丽娟,杨习贝,杨静宇,等. 基于覆盖的粗糙集模型比较[J]. 计算机科学,2012,39(7):229-232

[15] Wang Feng, Liang Ji-ye, Qian Yu-hua. Attribute reduction: A dimension incremental strategy, Knowledge-Based Systems, 2013, 39: 95-108

[16] Liang Ji-ye, Wang Feng, Dang Chuang-yin, et al. Incremental approach to feature selection based on rough set theory[J]. IEEE Transactions on Knowledge and Data Engineering, 2013

[17] Meng X. Posterior predictive values[J]. The Annals of Statistics, 1994(3): 1142-1160

[18] Ausin M C, Gomez-Villegas M A, Gonzalez-Perez B, et al. Bayesian Analysis of Multiple Hypothesis Testing with Applications to Microarray Experiments[J]. Communications in Statistics-Theory and Methods, 2011, 40(13): 2276-2291

[19] Li J D. Testing each hypothesis marginally at alpha while still controlling FWER: how and when[J]. Statistics in Medicine, 2012,32(10):1730-1738

[20] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing[J]. Journal of the Royal Statistical Society. Series B(Methodological), 1995, 57(1):289-300

[21] Qin W, Liu Y, Jiang T, et al. The Development of Visual Areas Depends Differently on Visual Experience[J]. PloS one, 2013, 8(1):e53784

[22] 刘晋, 张涛, 李康. 多重假设检验中 FDR 的控制与估计方法[J]. 中国卫生统计, 2012, 29(2):305-308

[23] Bilgin B, Brenner L. Context affects the interpretation of low but not high numerical probabilities: A hypothesis testing account of subjective probability[J]. Organizational Behavior and Human Decision Processes, 2013, 121(1):118-128

[24] Wang Y, Mei Y. A Multistage Procedure for Decentralized Sequential Multi-Hypothesis Testing Problems [J]. Sequential Analysis, 2012, 31(4):505-527

[25] 刘乐平, 张龙, 蔡正高. 多重假设检验及其在经济计量中的应用[J]. 统计研究, 2007, 24(4):26-30

[26] Yekutieli D, Benjamini Y. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics [J]. Journal of Statistical Planning and Inference, 1999, 82(1): 171-196

[27] Benjamini Y, Liu W. A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence[J]. Journal of Statistical Planning and Inference, 1999, 82(1):163-170

[28] Benjamini Y, Hochberg Y. On the adaptive control of the false discovery rate in multiple testing with independent statistics[J]. Journal of Educational and Behavioral Statistics, 2000, 25(1): 60-83

[29] Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency[J]. Annals of statistics, 2001, 29(4):1165-1188

[23] Benjamini Y, Krieger A, Yekutieli D. Adaptive linear step-up procedures that control the false discovery rate[J]. Biometrika, 2006, 93(3):491-507

[31] Benjamini Y. Discovering the false discovery rate[J]. Journal of the Royal Statistical Society: Series B(Statistical Methodology), 2010, 72(4):405-416

[32] Storey J. The positive false discovery rate: A bayesian interpretation and the q-value[J]. Annals of Statistics, 2003, 31(2003): 2013-2035

[33] Tusher V, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response[J]. Proceedings of the National Academy of Sciences, 2001, 98(9):5116-5121

[34] Efron B. Microarrays, empirical bayes and the two-groups model [J]. Statistical Science, 2008, 23(1):1-22

[35] Guedj M, Robin S, Celisse A, et al. Kerfdr: a semi-parametric kernel-based approach to local false discovery rate estimation [J]. BMC bioinformatics, 2009, 10(1):84-96

[36] Efron B. Large-scale simultaneous hypothesis testing [J]. Journal of the American Statistical Association, 2004, 99(465): 96-104

[37] Wu B, Guan Z, Zhao H. Parametric and nonparametric fdr estimation revisited[J]. Biometrics, 2006, 62(3):735-744

[38] Noble W. How does multiple testing correction work? [J]. Nature biotechnology, 2009, 27(12):1135-1137

[39] Xie Y, Pan W, Khodursky A. A note on using permutation-based false discovery rate estimates to compare different analysis methods for microarray data[J]. Bioinformatics, 2005, 21(23): 4280-4288

[40] Strimmer K. A unified approach to false discovery rate estimation[J]. BMC Bioinformatics, 2008, 9(1):303

[41] Muralidharan O. An empirical bayes mixture method for effect size and false discovery rate estimation[J]. The Annals of Applied Statistics, 2010, 4(1):422-438

[42] Kim K, Van De Wiel M. Effects of dependence in high-dimensional multiple testing problems[J]. BMC bioinformatics, 2008, 9(1):114

[43] Wille A, Zimmermann P, Vranová E, et al. Sparse graphical gaussian modeling of the isoprenoid gene network in arabidopsis thaliana[J]. Genome Biol, 2004, 5(11):R92

[44] Langaas M, Lindqvist B, Ferkingstad E. Estimating the proportion of true null hypotheses, with application to dna microarray data[J]. Journal of the Royal Statistical Society: Series B(Statistical Methodology), 2005, 67(4):555-572

[45] Benjamini Y. Discovering the false discovery rate[J]. Journal of the Royal Statistical Society: Series B(Statistical Methodology), 2010, 72(4):405-416

(上接第 69 页)

[2] 李斌, 王猛, 汪林, 等. 驾驶时间对营运驾驶员驾驶能力影响的试验研究[J]. 公路交通科技, 2007, 24(5):113-116

[3] 李都厚, 刘群, 袁伟, 等. 疲劳驾驶与交通事故关系 [J]. 交通运输工程学报, 2010, 10(2):104-109

[4] 毛吉吉, 初秀民, 严新平, 等. 汽车驾驶员驾驶疲劳监测技术研究进展 [J]. 中国安全科学学报, 2005, 15(3):108-112

[5] 戚基艳. 汽车驾驶疲劳分析及其监测[J]. 汽车科技, 2011(1): 34-38

[6] 宋义伟, 夏芹, 朱学峰. 驾驶员疲劳驾驶监测方法研究的进展 [J]. 自动化与信息工程, 2008, 28(4):31-34

[7] 冯舒. 基于驾驶座舱的驾车疲劳实验研究[D]. 合肥: 中国科学技术大学, 2007

[8] 瞿洋. 驾驶疲劳评测系统的设计[D]. 合肥: 中国科学技术大学, 2006

[9] 周传利. 司机疲劳监测系统中眼睛检测与跟踪研究[D]. 西安: 西安电子科技大学, 2008

[10] Thomy N, Thomas M N. Development of fatigue symptoms during simulated driving[J]. Accident Analysis and Prevention, 1999, 29:479-488

[11] 戚德虎, 康继昌. BP 神经网络的设计[J]. 计算机工程与设计, 1998, 19(2):48-50

[12] 杨凡, 赵建民, 朱信忠. 一种基于 BP 神经网络的车牌字符分类识别方法[J]. 计算机科学, 2006, 32(8):192-195