



Multi-class Depth-based Classification for Multivariate Data

Bayat, S. , Dehghan, S. 

Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran.

Corresponding author: S. Dehghan, sa_dehghan@sbu.ac.ir

Received: 14/1/2023 Revised: 13/8/2023 Accepted and Published Online: 17/8/2023.

Introduction

In a supervised classification problem with J classes, we have n_j labeled observations from the j -th class, $j = 1, \dots, J$. We use this training sample consisting of $n = \sum_{j=1}^J n_j$ observations to construct a decision rule for classifying an unlabeled observation \mathbf{x} to one of these J classes. Popular parametric classifiers like linear discriminant analysis and quadratic discriminant analysis are motivated by parametric model assumptions. So, they may lead to poor classification when these assumptions fail to hold. This paper presents a nonparametric multi-class depth-based classification approach for multivariate data. This approach is easy to implement rather than most existing nonparametric methods that have computational complexity. If the assumption of the elliptical symmetry holds, this method is equivalent to the Bayes optimal rule. The performance of the depth-based classifiers is evaluated using a Monte Carlo study as well as a real data example and is compared to that of several competitors.

Material and Methods

A depth function is aimed at ordering points in a space according to their centrality with respect to a distribution F . The larger the depth of a point \mathbf{x} , the more central \mathbf{x} is with respect to F . It is a useful tool in the multidimensional case where ordering the points from the inner to the outer part of a distribution or sample is not trivial task. Formally, a depth function is a bounded function that satisfies certain properties including affine invariance, maximised somewhere in the center of the distribution, monotonicity relative to deepest point, and vanishing at infinity. Therefore, the depth functions of-

fer a chance to exploit nonparametric methods in multivariate data analysis. Accordingly, several notions of multivariate depth have been proposed in the literature, and the field of applications of data depth is vast and still growing. This paper presents an overview of the depth function concept, and a non-parametric multi-class depth-based classification approach for multivariate data has been investigated. Moreover, the methodology has been verified using a simulation study and a real data set example. Calculations have been performed in R software, and R codes are available by request to the author.

Results and Discussion

The depth-based classification approach for multivariate data is equivalent to the Bayes optimal rule under elliptic symmetry. Indeed, under elliptic symmetry, the density function of a class can be expressed as a function of affine invariant depth functions, and hence the depth contours coincide with the density contours. Using both empirical and theoretical arguments, we consider the properties of the depth-based classification methods, and show that these classifiers can be particularly effective. We have used some simulated data sets to evaluate the performance of the depth-based classifiers. Moreover, to illustrate the method developed in this paper, we have applied the classifiers on a well-known data set.

Conclusion

In this article, the depth function and related classifiers have been investigated. We have analyzed several data sets simulated from elliptic as well as non-elliptic multivariate distributions. Moreover, the Swiss banknote data set has been analyzed for further evaluation of depth-based classifiers. The results showed that the maximum depth classifier based on the Mahalanobis and half-space depth functions has a good performance.

Keywords: Depth function, Classification, Bayes optimal rule, Elliptical symmetry.

Mathematics Subject Classification (2010): 62H30 62G99.



©The Author(s). The Publisher is Iranian Statistical Society.

This is an open access article distributed under the terms and conditions of [\(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/)

رده‌بندی چند-رده‌ای مبتنی بر تابع ژرفا برای داده‌های چندمتغیره

سارا بیات، سکینه دهقان

گروه آمار، دانشکده علوم ریاضی، دانشگاه شهید بهشتی

نویسنده مسئول: سکینه دهقان، sa_dehghan@sbu.ac.ir

تاریخ دریافت: ۱۴۰۱/۱۰/۲۴ تاریخ بازنگری: ۱۴۰۲/۵/۲۲ تاریخ پذیرش و انتشار: ۱۴۰۲/۵/۲۶

چکیده: این مقاله به بیان یک رویکرد ناپارامتری بر اساس تابع ژرفا برای رده‌بندی داده‌های چندمتغیره به چندین رده می‌پردازد. پیاده‌سازی این روش برخلاف اغلب روش‌های ناپارامتری دارای پیچیدگی محاسباتی نیست و در صورت برقراری فرض تقارن بیضوی مشاهدات، با قاعده بهینه بیزی معادل است. ارزیابی عملکرد این رده‌بندی‌ساز بر اساس توابع ژرفای مختلف، بر اساس مطالعات شبیه‌سازی و تحلیل داده‌های واقعی انجام می‌شود.

واژه‌های کلیدی: تابع ژرفا، رده‌بندی، تقارن بیضوی، قاعده بهینه بیزی
کد موضوع‌بندی ریاضی (۲۰۱۰): 62H30، 62G99.

۱ مقدمه

یکی از تمایزات بنیادین فنون یادگیری ماشین، بین روش‌های راهنماییده و ناراهنماییده است. زمانی از الگوریتم‌های یادگیری راهنماییده در رده‌بندی و پیشگویی استفاده می‌شود که داده‌هایی در دسترس است که در آن‌ها مقدار برآمد مورد نظر مشخص شده است. هنگامی که متغیر برآمدی برای پیشگویی یا رده‌بندی وجود ندارد، از الگوریتم‌های یادگیری ناراهنماییده استفاده می‌شود. قواعد پیوند، روش‌های کاهش بُعد و فنون خوشه‌بندی، همگی روش‌های ناراهنماییده هستند. به طور خاص، رده‌بندی که یک موضوع کاربردی در علم آمار است، روشی راهنماییده به منظور اختصاص یک مشاهده رده‌بندی نشده به یکی از دو یا چندین گروه یا جمعیت بر اساس اطلاعات تعدادی از متغیرها است. تجزیه و تحلیل چندمتغیره کلاسیک بر فرض نرمال چندمتغیره استوار است که از جمله این روش‌ها می‌توان به روش‌های بیزی،



تحلیل تشخیصی خطی درجه اول و درجه دوم و ... اشاره کرد. از طرفی در بسیاری از آزمایشات تصادفی، داده‌ها از توزیع نرمال پیروی نمی‌کنند. همچنین، برخی از رویکردهای پارامتری تحت تأثیر مشاهدات دورافتاده هستند؛ از این رو رویکردهای ناپارامتری برای رده‌بندی مشاهدات نسبت به رویکردهای پارامتری برتری دارند. از جمله رویکردهای ناپارامتری می‌توان به ماشین بردار تکیه‌گاه (وینیک، ۱۹۹۵)، قاعده K نزدیکترین همسایه (کاور و هارت، ۱۹۶۷) و قواعد رده‌بندی مبتنی بر توابع دورافتادگی رتبه‌ای (هوبرت و همکاران، ۲۰۱۷) اشاره کرد.

ژرفای داده ابزاری برای اندازه‌گیری میزان ژرفا^۱ بودن یا دور افتاده^۲ بودن یک نقطه معین نسبت به توده داده‌های چندمتغیره یا توزیع متناظر با آن است. هرچه یک نقطه نسبت به یک مجموعه داده و یا یک توزیع مرکزی‌تر باشد، دارای ژرفای بیش‌تری است و بالعکس میزان دورافتادگی آن کمتر است. به این ترتیب، راه‌های جدید و آسانی را برای تعیین بسیاری از ویژگی‌های پیچیده توزیع چندمتغیره، از جمله مکان، چندک‌ها، مقیاس، چولگی و کشیدگی ایجاد می‌کند. به بیان دقیق‌تر، با توجه به اینکه در روش‌های ناپارامتری یک‌متغیره، رتبه مشاهدات نقش بسیار تعیین‌کننده‌ای دارد، برای تعمیم این روش‌ها به حالت چندمتغیره، باید در این حالت نیز گونه‌ای رتبه‌بندی از مشاهدات امکان‌پذیر باشد. در حالت یک‌متغیره، ابتدا مشاهدات را مرتب کرده و سپس رتبه‌بندی می‌کنیم. تعمیم این روش به R^p برای $p \geq 2$ به دلیل فقدان چنین ترتیبی در R^p ، بی‌نتیجه است. برای جبران این نقیصه می‌توان ابتدا مرکزی برای داده‌های چندمتغیره تعیین کرد. این نقطه دارای بیشترین ژرفا خواهد بود و مابقی مشاهدات هرچه از مرکز در تمامی جهات دورتر باشند، ژرفای کمتری دارند. نقطه با بزرگترین مقدار ژرفا، بیشترین رتبه و به سایر نقاط به ترتیب بر اساس مقدار ژرفایشان رتبه اختصاص داده می‌شود. بنابراین بر اساس مقدار ژرفای نقاط، مشاهدات به صورت مرکز به بیرون رتبه‌بندی می‌شوند. روش‌های مختلفی برای به دست آوردن مرکز وجود دارد. تا به حال در حوزه‌های مختلف آمار بر اساس تابع ژرفا و رتبه‌بندی مبتنی بر آن پژوهش‌های بسیاری انجام شده است که از آن جمله می‌توان به شناسایی نقاط دورافتاده بر اساس تابع ژرفا از دهقان و فریدروhani (۲۰۲۱) اشاره کرد. در این مقاله، به دنبال معرفی روش رده‌بندی چندمتغیره مبتنی بر تابع ژرفا هستیم. بدین منظور، در بخش ۲، ابتدا مفهوم تابع ژرفا بیان شده و سپس چند تابع ژرفا مرور می‌شوند. در بخش ۳، قاعده رده‌بندی مبتنی بر تابع ژرفا مطرح می‌شود. مطالعات شبیه‌سازی و تحلیل داده‌های واقعی نیز به ترتیب در بخش‌های ۴ و ۵ بیان می‌شوند.

۲ تابع ژرفا

فرض کنید اندازه احتمال P بر فضای R^p داده شده و F تابع توزیع متناظر با آن باشد. ژرفای داده، اندازه‌ای از میزان ژرف بودن یا دورافتاده بودن یک نقطه نسبت به توده داده‌ها یا توزیع F است. هر تابع $D(\cdot, F)$ که ترتیبی مرکز به بیرون از نقاط x متعلق به R^p فراهم کند، تابع ژرفای متناظر با F نامیده می‌شود. بنابراین بر اساس تابع ژرفای مفروض، نقاط p متغیره به نحوی رتبه‌بندی می‌شوند که ژرفای بزرگتر نشان‌دهنده مرکزی‌تر بودن نقاط و ژرفای

¹Deep

²Outlier

کوچکتر نشان‌دهنده دورافتاده‌تر بودن آنها است. تحت تابع ژرفای مفروض، نقطه یا میانگین نقاط با بیشترین مقدار ژرفا، مرکز را تشکیل می‌دهند. در ادامه چند تابع ژرفا مرور می‌شوند.

ژرفای نیم‌فضا^۱ (توکی، ۱۹۷۵) نقطه $\mathbf{x} \in \mathbb{R}^p$ نسبت به توزیع F به صورت

$$HD(\mathbf{x}, F) = \inf \left\{ P(H) : H \text{ نیم‌فضایی بسته شامل } \mathbf{x} \text{ در فضای } p \text{ بعدی است} \right\}$$

تعریف می‌شود که در آن $P(H)$ احتمال در برداشتن نیم‌فضای H توسط توزیع F است. ژرفای نیم‌فضای نمونه‌ای نقطه \mathbf{x} نسبت به نمونه $\mathbf{X}_1, \dots, \mathbf{X}_n$ در \mathbb{R}^p عبارت است از کمترین تعداد نمونه تقسیم بر تعداد کل نمونه در بین تمامی نیم‌فضاهای بسته شامل نقطه \mathbf{x} .

ژرفای زونوئید^۲ (دایچرهوف و همکاران، ۱۹۹۶) نقطه $\mathbf{x} \in \mathbb{R}^p$ نسبت به توزیع F به صورت

$$ZD(\mathbf{x}, F) = \max \{ \alpha : \mathbf{x} \in D_\alpha(F) \}$$

تعریف می‌شود، که در آن $D_\alpha(F)$ ناحیه زونوئید نامیده شده، $D_\alpha(F) = R^d$ و برای $\alpha \in (0, 1]$ به صورت

$$D_\alpha(F) = \left\{ \int_{R^d} \mathbf{x}g(\mathbf{x})dF(\mathbf{x}) : 0 \leq g \leq \frac{1}{\alpha}, \int_{R^d} g(\mathbf{x})dF(\mathbf{x}) = 1 \right\}$$

تعریف می‌شود. تابع ژرفای نمونه‌ای زونوئید نیز با جایگذاری

$$D_\alpha(\mathbf{X}_1, \dots, \mathbf{X}_n) = \left\{ \sum_{i=1}^n \lambda_i \mathbf{x}_i : \sum_{i=1}^n \lambda_i = 1, 0 \leq \lambda_i \leq \frac{1}{\alpha}, i = 1, \dots, n \right\}$$

به جای $D_\alpha(F)$ در فرمول تابع ژرفای زونوئید حاصل می‌شود.

ژرفای ماهالانوبیس^۳ (لیو و سینگ، ۱۹۹۳) نقطه $\mathbf{x} \in \mathbb{R}^p$ نسبت به توزیع F به صورت

$$MD(\mathbf{x}, F) = \frac{1}{1 + d(\mathbf{x}, \boldsymbol{\mu}_F)} = \frac{1}{1 + (\mathbf{x} - \boldsymbol{\mu}_F)^T \boldsymbol{\Sigma}_F^{-1} (\mathbf{x} - \boldsymbol{\mu}_F)}$$

تعریف می‌شود، که در آن $\boldsymbol{\mu}_F$ و $\boldsymbol{\Sigma}_F$ به ترتیب بردار میانگین و ماتریس کوواریانس توزیع F و $d(\mathbf{x}, \boldsymbol{\mu}_F) = (\mathbf{x} - \boldsymbol{\mu}_F)^T \boldsymbol{\Sigma}_F^{-1} (\mathbf{x} - \boldsymbol{\mu}_F)$ فاصله ماهالانوبیس بین \mathbf{x} و $\boldsymbol{\mu}_F$ است. برای به دست آوردن نسخه نمونه‌ای ژرفای ماهالانوبیس نقطه $\mathbf{x} \in \mathbb{R}^p$ کافی است به جای $\boldsymbol{\mu}_F$ و $\boldsymbol{\Sigma}_F$ برآورد نمونه‌ای آن‌ها را جایگزین

^۱Halfspace Depth

^۲Zonoid Depth

^۳Mahalanobis Depth

کنیم، که به طور معمول به ترتیب از بردار میانگین نمونه‌ای و ماتریس کوواریانس نمونه‌ای استفاده می‌شود.

ژرفای تصویرمبنا^۱ (ژو، ۲۰۰۳) نقطه $\mathbf{x} \in \mathbb{R}^p$ نسبت به توزیع F به صورت

$$PD(\mathbf{x}, F) = \frac{1}{1 + \tilde{O}_P(\mathbf{x}, F)}$$

تعریف می‌شود که در آن $\tilde{O}_P(\mathbf{x}, F)$ به صورت

$$\tilde{O}_P(\mathbf{x}, F) = \sup_{\|\mathbf{u}\|=1} \left| \frac{\mathbf{u}^T \mathbf{x} - \mu(F_{\mathbf{u}^T \mathbf{X}})}{\sigma(F_{\mathbf{u}^T \mathbf{X}})} \right|$$

تعریف می‌شود که $\mu(\cdot)$ و $\sigma(\cdot)$ به ترتیب اندازه یک‌متغیره مکان و مقیاس توزیع F هستند، به گونه‌ای که هرگاه $\mathbf{u}^T \mathbf{x} - \mu(F_{\mathbf{u}^T \mathbf{X}}) = \sigma(F_{\mathbf{u}^T \mathbf{X}}) = 0$ ، عبارت $\left| \frac{\mathbf{u}^T \mathbf{x} - \mu(F_{\mathbf{u}^T \mathbf{X}})}{\sigma(F_{\mathbf{u}^T \mathbf{X}})} \right|$ مقدار صفر را اختیار می‌کند. برای به دست آوردن نسخه نمونه‌ای ژرفای تصویرپایه کافی است به جای توزیع F همتای نمونه‌ای آن، F_n جایگزین شود.

ژرفای فضایی^۲ (وردی و ژانگ، ۲۰۰۰) نقطه \mathbf{x} نسبت به توزیع F به صورت

$$SD(\mathbf{x}, F) = 1 - \|E(S(\mathbf{x} - \mathbf{X}))\|$$

تعریف می‌شود که در آن

$$S(\mathbf{x}) = \begin{cases} \frac{\mathbf{x}}{\|\mathbf{x}\|} & \mathbf{x} \neq 0, \\ 0 & \mathbf{x} = 0, \end{cases}$$

تابع علامت برداری در \mathbb{R}^p است. مقدار تابع فضایی نمونه‌ای نقطه \mathbf{x} به صورت

$$SD(\mathbf{x}, F_n) = 1 - \left\| \frac{1}{n} \sum_{i=1}^n S(\mathbf{x} - \mathbf{X}_i) \right\|$$

تعریف می‌شود.

¹Projection based Depth

²Spatial Depth

۳ قاعده رده‌بندی ژرفا مبنا

هدف هر قاعده رده‌بندی، یافتن قاعده یا ابزاری است که این امکان را فراهم کند که یک مشاهده را به یکی از چندین رده (یا کلاس) متفاوت اختصاص دهد. فرض کنید \mathbf{X} یک بردار تصادفی p -بعدی باشد که دارای یکی از J ، $J \geq 2$ ، توزیع مطلقاً پیوسته F_1, \dots, F_J در R^p است. یک رده‌بندی‌ساز فضای R^p را به J بخش مجزا A_1, \dots, A_J ، $\bigcup_{j=1}^J A_j = R^p$ ، تقسیم می‌کند، به طوری که \mathbf{x} به F_j اختصاص می‌یابد، اگر و تنها اگر $\mathbf{x} \in A_j$. به طور معادل

$$\mathbf{x} \in A_j \Leftrightarrow \text{class}(\mathbf{x}) = j.$$

فرض کنید f_j و π_j به ترتیب تابع چگالی و توزیع پیشین متناظر با F_j هستند، $j = 1, \dots, J$. قاعده بهینه بیزی، یک مشاهده را به رده یا توزیع با بیشترین احتمال پسین اختصاص می‌دهد. به عبارت دیگر \mathbf{x} به j امین کلاس اختصاص داده می‌شود، هرگاه $\pi_j f_j(\mathbf{x})$ بیش‌ترین باشد:

$$\text{class}_B(x) = \arg \max_{1 \leq j \leq J} \pi_j f_j(x). \quad (۱)$$

بر اساس این ایده و با توجه به مفهوم ژرفای داده، انتظار می‌رود که بر اساس تابع ژرفا نیز بتوان یک قاعده رده‌بندی تعریف است؛ زیرا ژرفای داده‌ها در مورد اینکه یک مشاهده چندمتغیره نسبت به توزیع مشخص F یا یک توده داده چندمتغیره چقدر دورافتاده است اطلاع می‌دهد. اگر یک مشاهده مرکزی باشد، مقدار ژرفای آن بزرگ و اگر یک مشاهده دورافتاده باشد، مقدار ژرفای آن کوچک خواهد بود. بنابراین ژرفای کوچک نشان دهنده دوری مشاهده مورد نظر نسبت به توزیع F است. **گوش و چادوری (۲۰۰۵)** یک قاعده رده‌بندی مبتنی بر تابع ژرفا برای اختصاص یک مشاهده جدید به هر یک از J رده متفاوت را به صورت

$$\text{class}_D(\mathbf{x}) = \arg \max_{1 \leq j \leq J} D(\mathbf{x}, F_j)$$

تعریف کردند. به بیان دقیق‌تر، ژرفای مشاهده مورد نظر نسبت به همه رده‌ها محاسبه می‌شود، سپس به رده‌ای با بیش‌ترین ژرفا اختصاص می‌یابد. بر اساس لم زیر نتیجه می‌شود، می‌توان تابع ژرفا را بر حسب تابع چگالی احتمال برای توزیع‌های متقارن بیضوی، بیان کرد.

لم ۱. (ژو و سرفلینگ، ۲۰۰۰) فرض کنید F_j توزیع متقارن بیضوی با پارامتر مکان θ_j و ماتریس کوواریانس Σ_j باشد. بنابراین تابع چگالی احتمال آن به صورت

$$f_j(\mathbf{x}) = |\Sigma_j|^{-\frac{1}{2}} h((\mathbf{x} - \theta_j)' \Sigma_j^{-1} (\mathbf{x} - \theta_j))$$

۳۲۴ رده‌بندی چند-رده‌ای مبتنی بر تابع ژرفا

است که در آن $h(\cdot)$ یک تابع اکیداً نزولی، پیوسته و اسکالر نامنفی است. اگر تابع ژرفا ناوردای آفین^۱ باشد، سپس تابع صعودی $\omega(\cdot)$ وجود خواهد داشت که $f_j(\mathbf{x}) = \omega(D(\mathbf{x}, F_j))$.

بر اساس این لم، گوش و چادوری (۲۰۰۵) نشان دادند رده‌بندی‌ساز بیشترین ژرفا تحت برخی شرایط رده‌بندی‌ساز بیزی است. در ادامه این قضیه بیان می‌شود.

قضیه ۰.۱ (گوش و چادوری، ۲۰۰۵) تحت فرض تقارن بیضوی توزیع‌های F_1, \dots, F_J با ماتریسهای کوواریانس برابر، چنانچه احتمالات پیشین برابر باشند، رده‌بندی‌ساز (۱)، رده‌بندی‌ساز بیزی است.

برهان: فرض کنید $\text{class}_D(\mathbf{x})$ و $\text{class}_B(\mathbf{x})$ به‌ترتیب رده‌بندی‌ساز ژرفا مبنا و بیزی برای مشاهده \mathbf{x} باشند. سپس

$$\text{class}_B(\mathbf{x}) = \arg \max_{1 \leq j \leq J} \pi_j f_j(\mathbf{x}) = \arg \max_{1 \leq j \leq J} \omega(D(\mathbf{x}, F_j)).$$

تساوی آخر از برابری $\pi_1 = \dots = \pi_J$ و لم ۱ نتیجه شده است. حال با توجه به صعودی بودن تابع $\omega(\cdot)$ ، یافتن مقدار j ماکسیمم‌کننده $\omega(D(\mathbf{x}, F_j))$ معادل یافتن مقدار j ماکسیمم‌کننده $D(\mathbf{x}, F_j)$ است. بنابراین نتیجه مورد نظر حاصل شد.

در عمل، به دلیل نامعلوم بودن تابع توزیع، تابع ژرفا نیز نامعلوم است؛ و باید از نسخه تجربی آن استفاده کرد. در این صورت نسخه تجربی قاعده رده‌بندی نیز بر اساس تابع ژرفای تجربی تعریف می‌شود. چنانچه تابع ژرفای تجربی در شرط سازگاری $\sup_{\mathbf{x}} |D(\mathbf{x}, F_n) - D(\mathbf{x}, F)| \rightarrow 0$ صدق کند، این جایگزینی توجیه‌پذیر است. خوشبختانه شرط سازگاری برای توابع ژرفای نیم‌فضا، سادگی و زونوئید برقرار است، (ژو، ۲۰۰۳). همچنین تحت متناهی بودن گشتاورهای مرتبه دوم برای تابع ژرفای ماهالانوبیس و تصویر-مبنا نیز برقرار است، (لیو و سینگ، ۱۹۹۳).

۴ مطالعات شبیه‌سازی

در این بخش عملکرد روش پیشنهادی بر اساس توابع ژرفای مختلف مورد ارزیابی قرار گرفته و با الگوریتم k نزدیکترین همسایه، K-NN، که یکی از الگوریتم‌های راهنمایی‌دهنده در یادگیری ماشین برای رده‌بندی داده‌ها است، مقایسه می‌شود. ایده رده‌بندی K-NN شامل دو مرحله است؛ نخست نزدیکترین همسایه‌ها را تعیین می‌کند و در مرحله دوم رده‌ها با استفاده از آن همسایگان تعیین می‌شوند. مزیت عمده روش K-NN نداشتن فرضیه‌های پارامتری است، اما دارای پیچیدگی‌های محاسباتی است، به طوری که یافتن همسایه‌ها در یک مجموعه داده حجیم

¹Affine invariant

بسیار زمان بر است. فرض کنید π_1 و π_2 توزیع‌های دومتغیره متقارن بیضوی به ترتیب با مرکز تقارن μ_1 و μ_2 و ماتریس کوواریانس Σ_1 و Σ_2 باشند. همچنین فرض کنید احتمالات پیشین برابر هستند. فرض کنید X_1, \dots, X_n یک نمونه تصادفی از π_1 و Y_1, \dots, Y_m یک نمونه تصادفی از π_2 باشند. بردارهای تصادفی جدید Z_1, \dots, Z_m از π_1 و Z_{m+1}, \dots, Z_{m+n} از π_2 تولید می‌شوند. اندازه نمونه n و m برابر با ۱۰۰ و تعداد تکرار شبیه‌سازی ۱۰۰۰ در نظر گرفته شده است. نمونه‌ها از توزیع نرمال دو متغیره و توزیع لاپلاس دو متغیره تولید می‌شوند و رده‌بندی‌سازها بر اساس توابع ژرفای نیم‌فضا، ماهالانویس، تصویر-مینا، فضایی و زنونید محاسبه شده‌اند. برای محاسبه توابع ژرفا از بسته آماری ddalpha در نرم افزار R استفاده شده است. برای مقایسه عملکرد رده‌بندی‌سازها از نرخ بدرده‌بندی^۱ و معیارهای حساسیت^۲ و مشخص‌سازی^۳ استفاده می‌شود. واضح است خطایی که ممکن است در هنگام رده‌بندی رخ دهد، اختصاص دادن مشاهده جدید به یک زیر جامعه است در حالی که در اصل به زیر جامعه دیگری تعلق داشته باشد. بنابراین احتمال بدرده‌بندی به صورت $\Delta = \sum_{j=1}^J \pi_j P(x \notin A_j | F_j)$ تعریف می‌شود.

معیارهای حساسیت و مشخص‌سازی از دیگر معیارهای تشخیص درستی رده‌بندی هستند. فرض کنید C_1 و C_2 رده ۱ و ۲ باشند. تعریف کنید $n_{1,1}$ تعداد اعضای C_1 که به درستی رده‌بندی شده‌اند، $n_{2,2}$ تعداد اعضای C_2 که به درستی رده‌بندی شده‌اند، $n_{1,2}$ تعداد اعضای C_1 که به نادرستی در C_2 رده‌بندی شده‌اند و $n_{2,1}$ تعداد اعضای C_2 که به نادرستی در C_1 رده‌بندی شده‌اند. فرض کنید رده‌بندی درست عضویت در رده C_1 ، از رده C_2 مهم‌تر باشد، در این صورت حساسیت یک رده‌بندی‌کننده، توانایی آن در تشخیص عضویت در رده مهم، اعضای C_1 است، که به صورت $\frac{n_{1,1}}{n_{1,1} + n_{1,2}}$ تعریف می‌شود. همچنین مشخص‌سازی یک رده‌بندی‌کننده، به صورت $\frac{n_{2,2}}{n_{2,2} + n_{2,1}}$ تعریف می‌شود. ادامه فرض کنید، C_1 رده‌ای باشد که در آن نمونه‌ها بر اساس μ_1 و Σ_1 و C_2 رده‌ای باشد که در آن نمونه‌ها بر اساس μ_2 و Σ_2 تولید می‌شوند.

در اولین سناریوی شبیه‌سازی به بررسی عملکرد روش‌های مورد نظر در صورت اختلاف مکان بین رده‌ها می‌پردازیم. μ_1 و μ_2 به ترتیب $(0, 0)^T$ و $(\delta, 0)^T$ و $\Sigma_1 = \Sigma_2 = I_2$ در نظر گرفته می‌شوند، که $\delta \in [-2, 2]$. متوسط نرخ بدرده‌بندی، معیار حساسیت و معیار مشخص‌سازی برای توزیع نرمال و لاپلاس دومتغیره به ترتیب در جدول ۱ نمایش داده شده است. ملاحظه می‌شود، با توجه به متوسط خطای بدرده‌بندی متناظر با هر یک از توزیع‌ها، متوسط خطای بدرده‌بندی در توزیع لاپلاس دومتغیره تقریباً کمتر از متوسط خطای بدرده‌بندی در توزیع نرمال دومتغیره است و معیار حساسیت و مشخص‌سازی در توزیع لاپلاس دومتغیره تقریباً بیشتر از معیار حساسیت و مشخص‌سازی در توزیع نرمال دومتغیره است. همچنین در هر دو توزیع ذکر شده، تابع ژرفای نیم‌فضا بهترین عملکرد را در بین سایر توابع ژرفای بررسی شده دارد؛ چرا که متوسط خطای بدرده‌بندی متناظر با آن در هر دو توزیع کمتر و معیار حساسیت و مشخص‌سازی متناظر با آن در هر دو توزیع بیشتر است. همچنین رده‌بندی‌سازهای مبتنی بر تابع ژرفا در مقابل روش K-NN متوسط خطای کمتر و معیار حساسیت و مشخص‌سازی بیشتری دارند.

¹Misclassification rate

²Sensitivity

³Specificity

جدول ۱. اختلاف مکان تحت توزیع‌های دومتغیره

توزیع	δ	HD	MD	PD	SD	ZD	KNN
نرخ بدرده‌بندی	۰	۰/۴۲۰	۰/۴۹۹	۰/۵۰۰	۰/۵۰۱	۰/۴۴۶	۰/۶۵۰
		۰/۲۹۶	۰/۳۱۴	۰/۳۲۲	۰/۲۶۰	۰/۳۰۶	۰/۴۵۰
		۰/۱۵۳	۰/۱۶۰	۰/۱۶۷	۰/۱۲۹	۰/۱۵۸	۰/۳۷۵
	۰	۰/۴۲۸	۰/۵۰۰	۰/۵۰۰	۰/۵۰۰	۰/۴۲۸	۰/۵۸۰
		۰/۲۳۵	۰/۲۵۷	۰/۲۶۰	۰/۳۱۶	۰/۲۴۹	۰/۴۰۰
		۰/۱۱۵	۰/۱۲۶	۰/۱۳۰	۰/۱۶۱	۰/۱۲۱	۰/۱۸۰
معیار حساسیت	۰	۰/۹۸۵	۰/۹۷۶	۰/۹۱۵	۰/۹۶۸	۰/۹۶۵	۰/۹۱۰
		۰/۹۵۵	۰/۹۷۲	۰/۹۷۳	۰/۹۶۰	۰/۹۲۸	۰/۹۲۸
		۰/۹۸۶	۰/۹۸۴	۰/۹۸۷	۰/۹۷۱	۰/۹۸۰	۰/۹۴۲
	۰	۰/۹۵۷	۰/۹۳۵	۰/۹۲۷	۰/۹۵۴	۰/۹۵۴	۰/۹۲۷
		۰/۹۷۵	۰/۹۵۳	۰/۹۶۱	۰/۹۷۱	۰/۹۷۸	۰/۹۴۲
		۰/۹۹۴	۰/۹۹۱	۰/۹۸۷	۰/۹۸۴	۰/۹۹۱	۰/۹۵۴
معیار مشخص‌سازی	۰	۰/۹۹۵	۰/۹۲۹	۰/۹۹۱	۰/۹۳۲	۰/۹۳۷	۰/۸۱۹
		۰/۹۸۷	۰/۹۷۳	۰/۹۵۸	۰/۹۸۱	۰/۹۷۲	۰/۸۳۷
		۰/۹۸۶	۰/۹۸۸	۰/۹۸۹	۰/۹۸۹	۰/۹۸۹	۰/۸۴۲
	۰	۰/۹۶۸	۰/۹۵۹	۰/۹۶۷	۰/۹۳۷	۰/۹۶۳	۰/۹۲۷
		۰/۹۹۴	۰/۹۸۲	۰/۹۸۱	۰/۹۷۰	۰/۹۷۴	۰/۹۴۲
		۰/۹۹۶	۰/۹۸۴	۰/۹۸۹	۰/۹۸۲	۰/۹۸۲	۰/۹۵۴

حال در سناریوی دوم عملکرد رده‌بندی‌ساز مبتنی بر تابع ژرفای نیم‌فضا را تحت اختلاف همزمان مکان و مقیاس بین رده‌ها بررسی می‌کنیم. نتایج در جدول ۲ ملاحظه می‌شود. μ_1 و μ_2 به ترتیب $(\circ, \circ)^T$ و $(\delta, \circ)^T$ که $\delta \in [-2, 2]$ و $\Sigma_1 = \mathbf{I}_2$ و $\Sigma_2 = 9\mathbf{I}_2$ در نظر گرفته می‌شوند. با توجه به اختلاف مقیاس در بین دو رده، همچنان تابع ژرفای نیم‌فضا عملکرد خوبی دارد و برای توزیع لاپلاس دومتغیره بهتر از توزیع نرمال دومتغیره عمل می‌کند.

حال در این بخش از شبیه‌سازی به بررسی عملکرد مجدد روش‌های مورد نظر در حالت چندمتغیره می‌پردازیم. بدین منظور، اندازه نمونه n و m برابر با ۱۰۰ و تعداد تکرار شبیه‌سازی ۱۰۰۰ در نظر گرفته شده است. نمونه‌ها از توزیع نرمال پنج متغیره و توزیع لاپلاس پنج متغیره تولید می‌شوند. μ_1 و μ_2 به ترتیب $(\circ, \circ, \circ, \circ)^T$ و $(\delta, \circ, \circ, \circ)^T$ و $\Sigma_1 = \Sigma_2 = \mathbf{I}_5$ در نظر گرفته می‌شوند، که $\delta \in [-2, 2]$. متوسط نرخ بدرده‌بندی، معیار حساسیت و معیار مشخص‌سازی برای توزیع نرمال و لاپلاس پنج‌متغیره به ترتیب در جداول ۳ نمایش داده شده است. برخلاف شبیه‌سازی از توزیع‌های دومتغیره، با توجه به متوسط خطای بدرده‌بندی متناظر با هر یک از توزیع‌ها، متوسط خطای بدرده‌بندی در توزیع نرمال پنج‌متغیره تقریباً کمتر از متوسط خطای بدرده‌بندی در توزیع لاپلاس پنج‌متغیره است.

جدول ۲. رده‌بندی ساز مبتنی بر تابع ژرفای نیم‌فضا برای اختلاف در مکان و مقیاس

		δ				
		۲	۱	۰	-۱	-۲
نرخ بدرده‌بندی	نرمال	۰/۳۸۰	۰/۴۵۹	۰/۴۹۰	۰/۴۶۰	۰/۳۸۰
	لاپلاس	۰/۲۹۹	۰/۴۰۶	۰/۴۸۴	۰/۴۰۹	۰/۲۹۹
معیار حساسیت	نرمال	۰/۹۸۷	۰/۹۶۴	۰/۹۴۸	۰/۹۷۹	۰/۹۷۹
	لاپلاس	۰/۹۸۷	۰/۹۸۲	۰/۹۴۸	۰/۹۷۱	۰/۹۸۹
معیار مشخص‌سازی	نرمال	۰/۹۸۴	۰/۹۸۳	۰/۹۵۴	۰/۹۶۳	۰/۹۸۸
	لاپلاس	۰/۹۸۹	۰/۹۷۸	۰/۹۷۱	۰/۹۷۸	۰/۹۹۳

ملاحظه می‌شود، معیار حساسیت در توزیع لاپلاس پنج‌متغیره تقریباً کمتر از معیار حساسیت در توزیع نرمال پنج‌متغیره است، در حالی که معیار مشخص‌سازی توزیع لاپلاس پنج‌متغیره تقریباً بیشتر از معیار حساسیت و مشخص‌سازی در توزیع نرمال پنج‌متغیره است. همچنین در هر دو توزیع ذکر شده، تابع ژرفای زونوئید بهترین عملکرد را در بین سایر توابع ژرفای بررسی شده دارد. همچنین رده‌بندی سازهای مبتنی بر تابع ژرفا در مقابل روش K-NN مجدد متوسط خطای کمتر و معیار حساسیت و مشخص‌سازی بیشتری دارند.

حال عملکرد رده‌بندی‌ساز مبتنی بر تابع ژرفای زونوئید را تحت اختلاف همزمان مکان و مقیاس بین رده‌ها بررسی می‌کنیم. نتایج در جدول ۴ ملاحظه می‌شود. μ_1 و μ_2 به ترتیب $(0, 0, 0, 0)^T$ و $(\delta, 0, 0, 0)^T$ که $\delta \in [-2, 2]$ و $\Sigma_1 = I_4$ و $\Sigma_2 = 9I_4$ در نظر گرفته می‌شوند. با توجه به اختلاف مقیاس در بین دو رده، همچنان تابع ژرفای زونوئید عملکرد خوبی دارد و برای توزیع لاپلاس پنج‌متغیره بهتر از توزیع نرمال پنج‌متغیره عمل می‌کند.

۵ رده‌بندی داده‌های اسکناس سوئیزی

در این بخش برای نشان دادن عملکرد روش رده‌بندی مبتنی بر تابع ژرفا یک مجموعه داده واقعی تحلیل می‌شود. داده‌ها شامل ۱۰۰ اسکناس اصلی و ۱۰۰ اسکناس جعلی ۱۰۰۰ فرانکی قدیمی سوئیس هستند؛ که شش متغیر طول اسکناس (برحسب میلی‌متر)، عرض سمت چپ اسکناس (برحسب میلی‌متر)، عرض سمت راست اسکناس (برحسب میلی‌متر)، عرض حاشیه پایین اسکناس (برحسب میلی‌متر)، عرض حاشیه بالا اسکناس (برحسب میلی‌متر) و قطر اسکناس (برحسب میلی‌متر) برای آن‌ها ثبت شده است. عملکرد رده‌بندی‌ساز پیشنهادی تحت توابع ژرفای مختلف برای رده‌بندی اسکناس‌ها به دو رده‌ی اسکناس اصل و اسکناس تقلبی بررسی شد. ابتدا فرض نرمال چندمتغیره، سپس فرض تقارن بیضوی مجموعه داده‌های شش متغیره را مورد بررسی قرار می‌دهیم. بسته‌های آماری "MVN" و "ellipticalsymmetry" در نرم افزار R به ترتیب شامل آزمون‌هایی برای بررسی نرمال چندمتغیره و تقارن بیضوی چندمتغیره مشاهدات هستند. بر اساس بسته آماری "MVN"، آماره‌های آزمون نرمال

جدول ۳. اختلاف مکان تحت توزیع‌های پنج‌متغیره نرمال و لاپلاس

توزیع	δ	HD	MD	PD	SD	ZD	KNN
نرخ بدرده‌بندی	نرمال	۰	۰/۲۶۴	۰/۴۹۸	۰/۵۰۱	۰/۲۵۸	۰/۵۹۳
	۱	۰/۲۴۴	۰/۲۵۸	۰/۱۴۴	۰/۱۵۱	۰/۰۸۲	
	۲	۰/۱۸۲	۰/۰۰۱	۰/۰۱۵	۰/۱۶۶	۰/۰۰۳	
لاپلاس	۰	۰/۳۳۵	۰/۴۹۹	۰/۵۰۱	۰/۵۰۲	۰/۳۳۰	۰/۶۹۳
	۱	۰/۲۶۳	۰/۰۵۹	۰/۱۲۲	۰/۱۲۴	۰/۰۶۳	
	۲	۰/۲۱۲	۰/۰۰۳	۰/۰۲۶	۰/۲۷۵	۰/۰۰۵	
معیار حساسیت	نرمال	۰	۰/۹۷۲	۰/۹۴۴	۰/۹۵۳	۰/۹۷۵	۰/۹۲۳
	۱	۰/۹۹۱	۰/۹۹۲	۰/۹۷۵	۰/۹۹۰	۰/۹۶۲	
	۲	۱	۰/۹۹۸	۱	۱	۱	۰/۹۹۱
لاپلاس	۰	۰/۹۵۴	۰/۹۳۸	۰/۹۹۳	۰/۹۶۶	۰/۹۷۵	۰/۹۲۶
	۱	۰/۹۸۳	۰/۹۸۰	۰/۹۷۶	۰/۹۸۳	۰/۹۶۲	
	۲	۰/۹۹۹	۰/۹۹۹	۰/۹۹۶	۰/۹۹۸	۱	۰/۹۷۳
معیار مشخص‌سازی	نرمال	۰	۰/۹۷۶	۰/۹۵۲	۰/۹۴۲	۰/۹۷۹	۰/۹۲۳
	۱	۰/۹۹۳	۰/۹۹۰	۰/۹۹۰	۰/۹۸۷	۰/۹۹۲	
	۲	۱	۱	۰/۹۹۷	۱	۱	۰/۹۸۶
لاپلاس	۰	۰/۹۸۱	۰/۹۵۶	۰/۹۷۴	۰/۹۴۲	۰/۹۵۳	۰/۹۳۰
	۱	۰/۹۹۵	۰/۹۹۲	۰/۹۹۱	۰/۹۹۴	۰/۹۹۱	
	۲	۰/۹۹۸	۰/۹۹۸	۱	۰/۹۹۷	۱	۰/۹۸۶

جدول ۴. رده‌بندی‌ساز مبتنی بر تابع ژرفای زونوئید برای اختلاف در مکان و مقیاس

δ					
توزیع پنج‌متغیره	-۲	-۱	۰	۱	۲
نرخ بدرده‌بندی	نرمال	۰/۲۹۷	۰/۴۸۴	۰/۵۰۰	۰/۴۸۲
	لاپلاس	۰/۲۰۷	۰/۳۹۳	۰/۴۹۱	۰/۳۹۲
معیار حساسیت	نرمال	۱	۰/۹۹۲	۰/۹۷۵	۰/۹۹۱
	لاپلاس	۰/۹۹۶	۰/۹۹۱	۰/۹۴۰	۰/۹۹۷
معیار مشخص‌سازی	نرمال	۱	۰/۹۹۲	۰/۹۸۴	۰/۹۹۴
	لاپلاس	۱	۰/۹۹۴	۰/۹۸۷	۰/۹۹۴

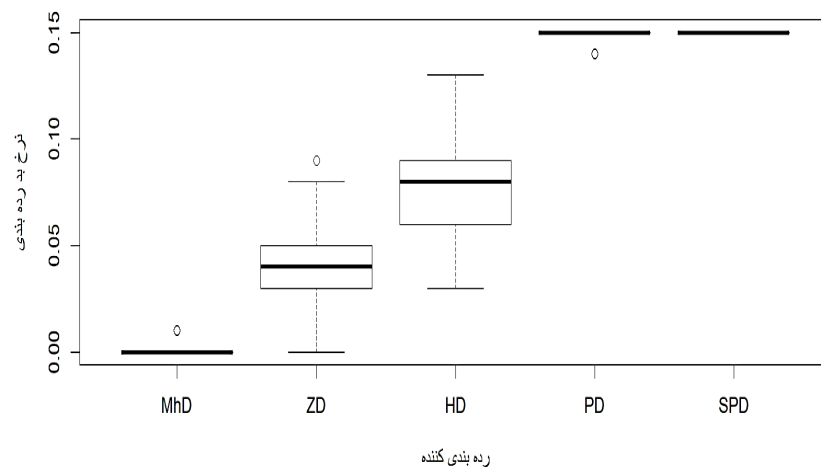
چندمتغیره از رویستون (۱۹۸۲)، هنز و زیرکالر (۱۹۹۰) و سیکلی و ریزو (۲۰۱۷) که به ترتیب با HZ_n ، R_n و SR_n نشان داده می شوند، و بر اساس بسته آماری "ellipticalsymmetry" آماره آزمون تقارن بیضوی بایبک و همکاران (۲۰۲۱) که با Ba_n نشان داده می شود، برای اسکناس های اصلی و جعلی محاسبه شده اند و نتایج در جدول ۵ آمده است. فرض نرمال چندمتغیره برای اسکناس های اصلی بر اساس آماره R_n در سطح ۰/۰۱ و

جدول ۵. نتایج آزمون های نرمال چندمتغیره و تقارن بیضوی چندمتغیره				
آزمون	اسکناس اصلی		اسکناس جعلی	
	آماره آزمون	p-value	آماره آزمون	p-value
R_n	۳۳/۶۵۸۲	$۵/۵ \times ۱۰^{-۶}$	۴۶/۶۶۳۶	$۱/۲ \times ۱۰^{-۸}$
HZ_n	۰۰/۸۸۱۶	۰/۶۹۱۵	۰۱/۳۰۰۳	$۲/۲ \times ۱۰^{-۱۱}$
SR_n	۰۰/۱۳۵۱	۰/۰۶۸۰	۰۲/۱۶۵۳	۰/۰۰۰۰
Ba_n	۱۰/۶۲۰۰	۰/۱۰۰۹	۴۴/۰۸۷۰	$۷/۱ \times ۱۰^{-۸}$

بر اساس آماره SR_n در سطح ۰/۱ رد می شود ولی بر اساس آماره های HZ_n و Ba_n به ترتیب دلیلی برای رد فرض نرمال چندمتغیره و فرض تقارن بیضوی داده های اسکناس های اصلی وجود ندارد. همچنین هر دو فرض مورد بررسی بر اساس همه آزمون ها برای اسکناس های جعلی در سطح ۰/۰۱ رد می شوند. بنابراین می توان نتیجه گرفت، قواعد رده بندی مبتنی بر فرضیات پارامتری برای این مجموعه داده قابل استفاده نیست. حال قواعد رده بندی مبتنی بر تابع ژرفا را برای رده بندی اسکناس های اصلی به یکی از رده های اصلی و جعلی به کار می بریم. رده بندی سازهای مبتنی بر توابع ژرفای نیم فضا، تصویر-مبنا و زنوئید به درستی همه ۱۰۰ اسکناس را در رده اسکناس های اصلی رده بندی می کنند. رده بندی سازهای مبتنی بر توابع ژرفای ماهالانوبیس و فضایی به درستی ۹۹ اسکناس را در رده اسکناس های اصلی رده بندی می کنند ولی اسکناس ۱۷۰م را به اشتباه به عنوان جعلی رده بندی می کنند. حال برای مقایسه دقیق تر رده بندی سازها، ۱۰۰۰ مرتبه به تصادف ۸۵ درصد مشاهدات را به عنوان داده های آموزشی و ۱۵ درصد مابقی به عنوان داده های آزمون انتخاب می شوند. متوسط نرخ بدرده بندی برای رده بندی ساز مبتنی بر تابع ژرفای نیم فضا، ماهالانوبیس، تصویر-مبنا، فضایی، زنوئید و روش $k-NN$ به ترتیب برابر با ۰/۷۹۲، ۰/۰۰۰۹، ۰/۱۴۹۰، ۰/۱۵۰۰، ۰/۴۱۱ و ۰/۱۹۶۰ است. ملاحظه می شود، رده بندی سازهای مبتنی بر تابع ژرفاهای ماهالانوبیس و زنوئید بهترین عملکرد را در بین سایر رده بندی سازها دارند.

بحث و نتیجه گیری

در این مقاله، روش رده بندی چندمتغیره ماکسیم ژرفا مورد بررسی قرار گرفت. پیاده سازی این روش نسبت به بیشتر روش های چندمتغیره که دارای پیچیدگی محاسباتی هستند، ساده تر است. بر اساس مطالعات داده های اسکناس



شکل ۱. نمودار جعبه‌ای نرخ بدرده‌بندی مرتبط با رده‌بندی‌سازها برای داده‌های اسکناس سوئیسی

سوئیسی، عملکرد رده‌بندی‌ساز ماکسیم ژرفا بر اساس توابع ژرفای مختلف مورد بررسی قرار گرفت. نتایج نشان داد، رده‌بندی‌ساز ماکسیم ژرفا بر اساس تابع ژرفای ماهالانوبیس و نیم‌فضا دارای عملکرد مطلوب‌تری است.

تقدیر و تشکر

نویسندگان مقاله از سردبیر، داوران و ویراستار محترم مجله که نظرات ارزشمندشان باعث بهبود مطالب ارائه شده در این مقاله گردید، کمال تشکر و قدردانی را دارد.

مراجع

دهقان، س. و فریدروحانی، م. (۱۴۰۰)، شناسایی نقاط دورافتاده چندمتغیره بر اساس تابع دورافتادگی ژرفا-مبنا، مجله علوم آماری، ۱۵(۲)، ۴۴۳-۴۶۲.

Babić, S., Gelbgras, L., Hallin, M., & Ley, C. (2021), Optimal Tests for Elliptical Symmetry: Specified and Unspecified Location, Bernoulli, 27(4), 2189-2216.

- Cover, T. M., Hart, P. E. (1967), Nearest Neighbor Pattern Classification, *IEEE Transactions on Information Theory*, **13**(1), 21–27.
- Dyckerhoff, R., and Mosler, K., and Koshevoy, G. (1996), Zonoid Data Depth: Theory and Computation. *Proceedings in Computational Statistics 12th Symposium held in Barcelona*, 235-240.
- Ghosh, A. K., and Chaudhuri, P. (2005), On maximum Depth and Related Classifiers, *Scandinavian Journal of Statistics*, **32**(2), 327-350.
- Henze, N. and Zirkler, B. (1990), A Class of Invariant Consistent Tests for Multivariate Normality. *Communications in Statistics - Theory and Methods*, **19**(10): 3595-3618.
- Hubert, M., Rousseeuw, P., and Segaeert, P. (2017), Multivariate and Functional Classification Using Depth and Distance, *Advances in Data Analysis and Classification*, **11**(3), 445-466.
- Liu, R. Y., and Singh, K. (1993), A quality Index Based on Data Depth and Multivariate Rank Tests, *Journal of the American Statistical Association*, **88**(421), 252–260.
- Royston, J.P. (1982), An Extension of Shapiro and Wilks W Test for Normality to Large Samples, *Applied Statistics*, **31**(2), 115-124.
- Székely, G. J., and Rizzo, M. L. (2017), The Energy of Data, *Annual Review of Statistics and Its Application*, **4**, 447-479.
- Tukey, J. W. (1975), Mathematics and the Picturing of Data, *Proceedings of the International Congress of Mathematicians*, **2**, 523–531.
- Vardi, Y. and Zhang, C. H. (2000), The Multivariate L1-median and Associated Data Depth, *Proceedings of the National Academy of Sciences*, **97**(4), 1423-1426.
- Vapnik, V. (1995), *The Nature of Statistical Learning Theory*, Springer-Verlag.

مراجع	٣٣٢
-------------	-----

Zuo, Y. (2003), Projection-based Depth Functions and Associated Medians, *The Annals of Statistics*, **31**(5), 1460–1490.

Zuo, Y., and Serfling, R. J. (2000), Structural Properties and Convergence Results for Contours of Sample Statistical Depth Functions, *Annals of Statistics*, **28**(2), 483-499.