

## *The False Discovery Rate in Simultaneous Fisher and Adjusted Permutation Hypothesis Testing on Microarray Data*

Mohsen Salehi<sup>1\*</sup> , Isaac Almasi<sup>2</sup> 

<sup>1</sup>Department of Statistics,  
School of Basic Sciences,  
University of Qom, Qom,  
Iran.

<sup>2</sup>Department of Mathematics,  
School of Basic Sciences,  
University of Ilam, Ilam,  
Iran.

\*Corresponding Author:  
**Mohsen Salehi**; Department  
of Statistics, School of Basic  
Sciences, University of Qom,  
Qom, Iran.

Email:  
m.salehi@qom.ac.ir

Received: 30 Jan, 2019  
Accepted: 1 Jul, 2019

### **Abstract**

**Background and Objectives:** In recent years, new technologies have led to produce a large amount of data and in the field of biology, microarray technology has also dramatically developed. Meanwhile, the Fisher test is used to compare the control group with two or more experimental groups and also to detect the differentially expressed genes. In this study, the false discovery rate was investigated in the simultaneous Fisher and adjusted permutation hypothesis testing on microarray data.

**Methods:** In this study, first, false discovery rate was computed through the simulation study and selection of three different modes for the samples of control group and experimental groups, then, these two methods, were applied to 8799 genes related to brain cell of the 29 rats (in three age groups of young, middle-aged, and aged), and the effect of the process of brain aging, was investigated on increased development of the Alzheimer disease.

**Results:** The results showed that the Fisher permutation methods cannot control the false discovery rate, but the adjusted permutation method works better and detects the significant differences more accurately, therefore, the number of false positives decrease in the second method.

**Conclusion:** Considering the results of this study, use of the customary methods such as the Fisher permutation test, which is the base of analyzing biological data in many of the software, do not have suitable efficiency in the large-scale data, including microarray data, and cannot control the false discovery rate, whereas the justified permutation method with better performance in false discovery rate leads to more reliable results.

**Keywords:** Microarray data; Gene expression; Alzheimer disease; False discovery rate; Methods.

DOI: 10.29252/qums.13.6.47

## نرخ کشف کاذب در آزمون فرض جایگشتی فیشر و جایگشتی تعدیل یافته همزمان بر داده‌های ریزآرایه

محسن صالحی\*<sup>۱</sup> ID، اسحاق الماسی<sup>۲</sup> ID

### چکیده

**زمینه و هدف:** در سال‌های اخیر، فن آوری‌های جدید منجر به تولید حجم انبوهی از داده‌ها شده و در حوزه زیستی، فناوری ریزآرایه نیز به صورت چشمگیری توسعه یافته است. در این میان، جهت مقایسه گروه کنترل با دو یا چند گروه آزمایشی، همچنین یافتن ژن‌هایی با بیان متفاوت، از آزمون فیشر استفاده می‌شود. در این مطالعه نرخ کشف کاذب در آزمون فرض جایگشتی فیشر و جایگشتی تعدیل یافته همزمان بر داده‌های ریزآرایه بررسی گردید.

**روش بررسی:** در این مطالعه، ابتدا با شبیه‌سازی و انتخاب سه حالت مختلف برای نمونه‌های کنترل و آزمایشی، نرخ کشف کاذب با استفاده از دو روش آزمون جایگشتی فیشر و جایگشتی تعدیل یافته محاسبه گردید، سپس این دو روش بر روی ۸۷۹۹ ژن مربوط به سلول مغز ۲۹ موش (در سه گروه سنی جوان، میانسال و پیر) اعمال گردید و تأثیر فرآیند سن مغز بر افزایش ایجاد بیماری آلزایمر مورد بررسی قرار گرفت.

**یافته‌ها:** نتایج نشان داد استفاده از روش آزمون جایگشتی فیشر، نرخ کشف کاذب را نمی‌تواند کنترل کند، ولی روش جایگشتی تعدیل یافته بهتر عمل کرده و اختلافات واقعی معنی‌دار را درست‌تر تشخیص می‌دهد؛ لذا مقدار FP در روش دوم کاهش خواهد یافت.

**نتیجه‌گیری:** با توجه به نتایج این مطالعه، استفاده از روش‌های مرسوم از جمله آزمون جایگشتی فیشر که مبنای تحلیل داده‌های زیستی در بسیاری از نرم‌افزارها می‌باشد، در داده‌های بزرگ مقیاس، از جمله داده‌های ریزآرایه کارایی مطلوب را ندارد و نرخ کشف کاذب را نمی‌تواند کنترل کند؛ درحالی‌که روش جایگشتی تعدیل یافته با عملکرد بهتر در کنترل نرخ کشف کاذب، نتایج قابل اعتمادتری در پی دارد.

**کلیدواژه‌ها:** ریزآرایه؛ بیان ژن؛ بیماری آلزایمر؛ نرخ کشف کاذب؛ روش‌ها.

<sup>۱</sup>گروه آمار، دانشکده علوم پایه، دانشگاه قم، قم، ایران.

<sup>۲</sup>گروه ریاضی، دانشکده علوم پایه، دانشگاه ایلام، ایلام، ایران.

\*نویسنده مسئول مکاتبات:

محسن صالحی؛ گروه آمار، دانشکده علوم پایه، دانشگاه قم، قم، ایران.

آدرس پست الکترونیکی:

m.salehi@qom.ac.ir

تاریخ دریافت: ۹۷/۱۱/۱۰

تاریخ پذیرش: ۹۸/۴/۱۰

لطفاً به این مقاله به صورت زیر استناد نمایید:

Salehi M, Almasi I. The false discovery rate in simultaneous fisher and adjusted permutation hypothesis testing on microarray data. Qom Univ Med Sci J 2019;13(6):47-54. [Full Text in Persian]

## مقدمه

قابل ملاحظه‌ای به دست آورد (۲). Tibshirani, Estorey, Efron. Xie, Pan و Zhao و همکاران روش جایگشتی را به‌طور وسیعی مورد استفاده قرار دادند (۳-۵). در همین راستا، Cui, Xu, Gao، Pesarin و همکاران نیز روش‌های معادل جایگشتی را به کار بردند (۶-۱۰). همچنین Bensal و همکاران نشان دادند زمانی که توزیع متغیرهای تصادفی نامتقارن باشد، آزمون فرض چندگانه مطرح است (۱۱). صالحی و همکاران با ارائه آماره صفر دیگری توانستند  $p$ -مقدار را دقیق‌تر برآورد کنند (۱۲).

نکته قابل تأمل این است که در برخی موارد محققین با بیش از دو جامعه یا دو گروه مستقل روبرو می‌شوند؛ به‌عنوان مثال، Blalock برای بررسی تغییر بیان ژن‌ها در بیماری آلزایمر، داده‌های ریز آرایه‌ای مربوط به سلول مغز یک نوع موش را در سه گروه سنی (جوان، میانسال و پیر) گردآوری و میزان بیان ۸۷۹۹ ژن را برای آن‌ها اندازه‌گیری کرد (۱۳)؛ بنابراین لزوم روش‌های مشابه با آنچه در مورد دو گروه مستقل بیان شد، در این داده‌های بزرگ مقیاس نیز احساس می‌شود. در این راستا تحقیقاتی توسط Zhao و Die، Blalock و همکاران نیز صورت گرفت (۱۳-۱۵). Blalock، یک غربالگری سه مرحله‌ای را برای کاهش تعداد مقایسه‌ها اعمال کرد، سپس با استفاده از آزمون فیشر کلاسیک، ژن‌هایی را که میزان بیان ژن در آن‌ها به‌صورت معنی‌داری تغییر کرده بود، گزارش کرد. در ادامه صالحی و همکاران، آماره تعدیل یافته جایگشتی دیگری برای مقایسه چند گروه مستقل ارائه دادند و تعمیمی از آن را مطرح کردند (۱۶). صالحی و همکاران ضمن محاسبه مقدار  $p$ - در دو روش جایگشتی مرسوم فیشر و روش جایگشتی تعدیل یافته، نرخ کشف کاذب (FDR) را در آن‌ها مورد مقایسه قرار دادند و عملکرد این دو روش را در شناسایی درست ژن‌های تأثیرگذار بررسی کردند.

## روش بررسی

در این مطالعه، میزان شناسایی و کشف ژن‌های مؤثر در بروز یک بیماری، هم از طریق شبیه‌سازی با نرم‌افزار R و هم با استفاده از داده واقعی بررسی شد.

در سال‌های اخیر، فن آوری‌های جدید منجر به تولید حجم انبوهی از داده‌ها شده است. در این میان و در حوزه زیستی، فناوری ریز آرایه به‌صورت چشمگیری توسعه یافته است. این فناوری بستری را به‌وجود آورده تا میزان بیان ژن برای هزاران ژن به‌طور همزمان و در شرایط مختلف اندازه‌گیری شود؛ به‌عنوان مثال در مقایسه سلول‌های سرطانی و غیرسرطانی، اگر تفاوت میزان بیان ژن برای یک ژن خاص معنی‌دار باشد، می‌توان نتیجه گرفت آن ژن در ایجاد آن سرطان دخیل بوده است. با این مختصر توضیح، می‌توان فهمید چالش اصلی در تحلیل این داده‌ها، حجم وسیعی از داده‌ها بوده است؛ بنابراین برای هریک از هزاران ژن موجود در این داده‌ها، تحلیل آماری مناسبی نیاز است تا بتوان تشخیص داد که آیا میزان بیان ژن در شرایط ذکر شده تغییر کرده است یا خیر؟ در ظاهر، مسئله به‌صورت یک آزمون فرض ساده است که در آن فرض صفر، فرض برابری میانگین بیان ژن در این دو گروه می‌باشد. شاید در این حالت انجام یک آزمون  $Z$  دو نمونه‌ای مستقل برای هر ژن، کافی به نظر آید؛ اما از این نکته نباید غافل شد که در داده‌هایی همچون داده‌های ریز آرایه، این نظریه وجود دارد که تعداد زیادی از این ژن‌ها بر ایجاد سرطان دخیل نیستند و به همین دلیل استفاده مستقیم از روش‌های مرسوم باعث خارج شدن خطا از کنترل می‌شود؛ لذا نکته کلیدی این است که چگونه می‌توان آماره مرجعی مثل توزیع صفر برای این آماره  $Z$  تعیین کرد تا براساس آن، مقدار بحرانی را بتوان به‌گونه‌ای محاسبه کرد که اگر تعدادی ژن در مجموعه‌ای از ژن‌ها دارای مقدار  $Z$  بزرگتر از مقدار بحرانی باشند، بتوان ادعا نمود آن تعداد ژن، معنی‌دار بوده است. بر این مبنا، هرچه این توزیع صفر دقیق‌تر باشد، نتیجه بهتری حاصل خواهد شد. در این راستا، اغلب از آزمون‌های کلاسیک و رایجی مانند آزمون تی و فیشر استفاده می‌شود که این آزمون‌ها محافظه‌کار هستند. به‌همین دلیل Efron توزیع صفر مرجعی را از آماره  $Z$  ساخته که توزیع آن، توزیع صفر آماره  $Z$  را بهتر تقریب می‌زند و عملکرد آزمون را بهتر می‌کند (۱). Pan آماره‌ای بر پایه روش جایگشتی از توزیع دو نمونه‌ای مستقل ارائه داد و این متد را در روش‌های تحلیل معنی‌داری ریز آرایه و روش مدل آمیخته اعمال و نتایج

آزمون دیگری با فرمول:

$$F^* = \frac{\sum_{i=1}^k n'_i (\bar{X}_{i(1)} - \bar{X}_{i(2)})^2 / (k-1)}{\sum_{i=1}^k \sum_{l=1}^2 (n_{il} - 1) S_{i(l)}^2 / (N - 2k)} \sim F_{k-1, N-2k}$$

استفاده می‌شود که در آن هر نمونه به حجم  $n_i$  پس از

جایگشت دادن به دو زیر نمونه به حجم‌های  $n_{i1}$  و  $n_{i2}$  تقسیم

می‌شود و برای هر کدام از زیر نمونه‌ها میانگین‌ها  $\{\bar{X}_{i(1)}\}$  و  $\{\bar{X}_{i(2)}\}$  و واریانس‌ها:  $\{S_{i(1)}^2\}$  و  $\{S_{i(2)}^2\}$  محاسبه می‌گردد. در ضمن در این آماره مقدار  $n'_i$  با استفاده از فرمول:

$$\frac{1}{n'_i} = \frac{1}{n_{i1}} + \frac{1}{n_{i2}}$$

به دست آمد.

در هر کدام از روش‌ها، مقدار  $p$ ، نسبت تعداد دفعاتی است که مقدار آماره‌های آزمون مطرح شده پس از جایگشت در هر تکرار، بیشتر از مقدار اولیه آماره‌های فوق و پیش از جایگشت باشد (۲، ۱۰).

اگر  $FP$  فرض صفر درستی باشد که به غلط رد شده و  $TP$  فرض صفر نادرستی باشد که به غلط رد نشده است؛ بنابراین

$$FDR = \frac{FP}{FP + TP}$$

برای انجام آزمون فرض، دو مقدار  $0.03$  و  $0.05$  برای خطای اسمی  $\alpha$  و تعداد جایگشت برای رسیدن به مقدار  $p$  مناسب،  $1000$  جایگشت در نظر گرفته شد. این شبیه‌سازی  $1000$  مرتبه تکرار شد و در هر بار تکرار مقادیر  $FP$  و  $TP$  اندازه‌گیری شدند. برای هر مقدار خطای اسمی در نظر گرفته شده، از معدل رند شده مقادیر  $FP$  و  $TP$  در این  $1000$  بار تکرار، جهت تعیین مقدار  $FDR$  استفاده شد. در عمل ممکن است حجم نمونه در گروه کنترل و گروه‌های آزمایشگاهی به صورت‌های متفاوتی در نظر گرفته شود. برای بررسی کارایی این روش در حجم‌های مختلف این نمونه‌ها نیز سه حالت مختلف برای حجم نمونه گروه‌های کنترل و آزمایشگاهی مدنظر قرار گرفت: (۱). حالت اول که در آن تعداد نمونه در هر سه گروه کنترل و آزمایشگاهی برابر باشد (۶، ۶، ۶)؛ (۲)؛ حالت دوم که در آن تعداد نمونه در دو

برای شبیه‌سازی داده‌های زیستی، حالت‌های مختلفی برای گروه کنترل و تعدادی گروه آزمایشگاهی در نظر گرفته شد که در این مطالعه یک گروه کنترل و دو گروه آزمایشگاهی مدنظر قرار گرفت. حالت‌های دیگری نیز برای چند گروه آزمایشگاهی انجام شد که نتایج مشابهی رقم زد. به این منظور،  $1000$  ژن به نحوی تولید شد که  $900$  ژن تحت فرض برابری میزان بیان ژن (فرض صفر در آزمون فرض) و  $100$  ژن با فرض عدم برابری میزان بیان ژن (فرض مقابل در آزمون فرض) تولید شد. برای تولید ژن‌هایی با میزان بیان یکسان نیز نمونه‌های تصادفی برای گروه کنترل و گروه‌های آزمایشگاهی از توزیع نرمال با میانگین صفر و واریانس یک یعنی  $N(0, 1)$  تولید شد.

در مقابل، برای تولید ژن‌هایی با میزان بیان متفاوت، نمونه‌های تصادفی برای گروه کنترل و دو گروه آزمایشگاهی از توزیع‌های نرمال با میانگین‌های به ترتیب  $\mu_1, \mu_2, \mu_3$  و با واریانس

یکسان یک، یعنی  $N(\mu_1, 1)$ ،  $N(\mu_2, 1)$  و  $N(\mu_3, 1)$

تولید شدند. برای تولید ژن‌هایی با میزان بیان نابرابر نیز خود این میانگین‌ها از توزیع نرمال با میانگین صفر و واریانس  $4$ ، یعنی  $N(0, 4)$  به دست آمدند (۱۵). برای یافتن ژن‌هایی با میزان بیان

ژن نابرابر، از آماره‌های آزمون متناسب با دو روش جایگشتی مرسوم و جایگشتی تعدیل یافته استفاده شد.

آماره آزمون روش جایگشتی مرسوم، شبیه به آماره آزمون فیشر است

آزمون فیشر از آماره‌ای با فرمول:

$$F = \frac{\sum_{i=1}^k n_i (\bar{X}_{i.} - \bar{X})^2 / (k-1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i.})^2 / (N-k)} \Big|_{H_0} \sim F_{k-1, N-k}$$

استفاده می‌کند که در آن  $K$  و  $N$  به ترتیب تعداد گروه‌های

درگیر (گروه‌های کنترل و آزمایشگاهی) و تعداد کل نمونه در تمام گروه‌ها می‌باشد. همچنین  $\bar{X}_{i.}$  مقدار میانگین نمونه جایگشت داده شده در هر تکرار آزمایش و  $\bar{X}$ ، میانگین کل نمونه است؛ در حالی که در روش جایگشتی تعدیل یافته از آماره

در هر سه نمودار روش PP، مقدار نرخ کشف کاذب را نزدیکتر به مقدار اسمی خطای لحاظ شده، (۰/۰۵) که در نمودار با خطی موازی محور افقی مشخص است) تخمین زده است. در ضمن تغییرات این مقادیر، در روش PP کمتر از روش UP بوده و با افزایش حجم نمونه‌ها در گروه کنترل و گروه‌های آزمایشگاهی، این تغییرات کمتر نیز شده‌اند. اعمال این روش در سطح خطای ۰/۰۳ نیز نتایج مشابهی ایجاد کرده است. مقدار میانگین نرخ کشف کاذب برای این ۱۰۰۰ تکرار، برای روش‌های PP و UP، در دو سطح خطای اسمی ۰/۰۵ و ۰/۰۳ در جدول شماره ۱ جمع‌آوری شد که تا نزدیکترین رقم صحیح نیز گرد شدند. در این جدول، زمانی که خطای اسمی ۰/۰۵ و حجم نمونه برای گروه‌های کنترل، همچنین دو گروه آزمایشگاهی ۶ در نظر گرفته شد، مقدار نرخ کشف کاذب برای دو روش PP و UP به ترتیب ۰/۰۶۷ و ۰/۰۹۹ به دست آمد که نشان می‌دهد مقدار نرخ کشف کاذب در روش PP نسبت به روش UP، به مقدار اسمی خطای در نظر گرفته شده ۰/۰۵ نزدیکتر است. در همین سطح خطای اسمی و برای نمونه‌های ۷، ۱۰، ۱۰ (به ترتیب مربوط به گروه کنترل و دو گروه آزمایشگاهی)، مقدار نرخ کشف خطا برای دو روش PP و UP به ترتیب ۰/۰۶۵ و ۰/۰۹۷ به دست آمد که نشان می‌داد روش PP در مقایسه با روش UP، خطا را دقیق‌تر کنترل می‌کند. در سطح خطای اسمی، برای نمونه‌های ۷، ۱۰، ۱۲ در گروه کنترل و دو گروه آزمایشگاهی، نتایج مشابهی دیده شد که نشانگر عملکرد بهتر روش PP نسبت به روش UP بود. در سطح خطای اسمی ۰/۰۳ نیز مشاهدات و محاسبات جدول شماره ۱ نشان داد نرخ کشف کاذب در روش PP کمتر از روش UP بوده و خطا را دقیق‌تر کنترل می‌کند.

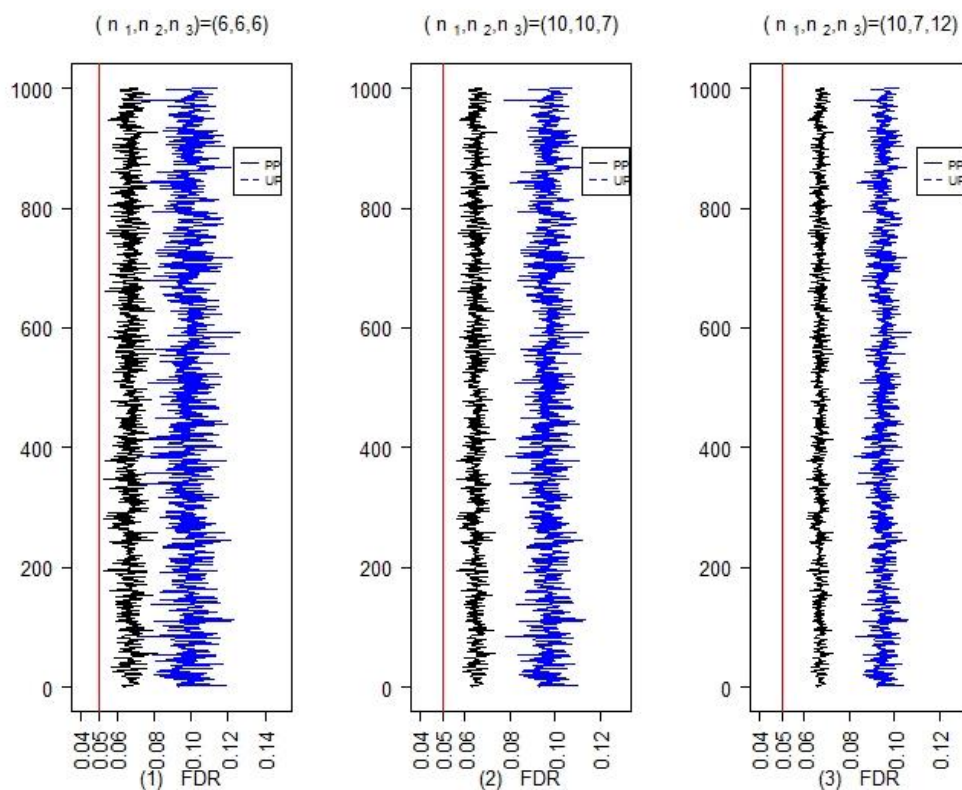
گروه آزمایشگاهی برابر باشد (۷، ۱۰، ۱۰، ۳)؛ حالت سوم که در آن تعداد نمونه در سه گروه مختلف متفاوت باشد (۷، ۱۰، ۱۲).

پس از آنکه مقادیر مشاهدات آزمون بر روی دو روش جایگشتی مرسوم و روش جایگشتی تعدیل یافته به دست آمد، مقادیر FP و TP برای دو روش یاد شده محاسبه و مقدار نرخ کشف کاذب در هر دو روش اندازه‌گیری شد.

داده‌های واقعی در این مطالعه مربوط به فرآیند سن مغز و تأثیر آن بر بیماری آلزایمر بود که توسط Blalock و همکاران (۵) ارائه گردید. در این مطالعه از نوعی موش نر در سه گروه سنی (جوان، میانسال و پیر) استفاده شد. سلول مغز این موش‌ها پس از استخراج در یک محیط پاک و طی فرآیندی به وسیله فناوری ریز آرایه بررسی و میزان بیان هزاران ژن در آن‌ها اندازه‌گیری شد. در این آزمایش، ۸۷۹۹ ژن شناسایی و میزان بیان ژن در آن‌ها به ترتیب برای ۹، ۱۰ و ۱۰ نمونه در هر گروه سنی اندازه‌گیری شد. در مطالعه حاضر از نرم‌افزار RMAexpress (برای نرمال‌سازی داده‌ها) استفاده شد. در نهایت، دو روش جایگشتی تعدیل یافته و روش جایگشتی مرسوم روی این داده‌ها اعمال گردید.

## یافته‌ها

نتایج شبیه‌سازی در شکل و جدول شماره ۱ آمده است. در شکل و جدول شماره ۱، UP نمادی برای روش جایگشتی مرسوم و PP نمادی برای روش جایگشتی تعدیل یافته است. در شکل مقادیر FDR برای ۱۰۰۰ تکرار در این آزمایش و در سطح خطای اسمی، ۰/۰۵ نشان داده شده است. همچنین، نمودارهای ۱، ۲ و ۳ به ترتیب مقادیر FDR را برای نمونه‌های مختلف در گروه‌های کنترل و آزمایشگاهی (۶، ۶، ۶)، (۷، ۱۰، ۱۰) و (۷، ۱۰، ۱۲) نمایش داده‌اند.



شکل: مقایسه مقادیر FDR در دو روش PP و UP با ۱۰۰۰ بار تکرار و به تفکیک نمونه‌های (۶،۶،۶)، (۱۰،۱۰،۷) و (۱۲،۱۰،۷) در گروه کنترل و دو گروه آزمایشگاهی. خطای اسمی ۰/۰۵ است.

جدول شماره ۱: میانگین نرخ کشف کاذب در دو روش جایگشتی مرسوم و روش جایگشتی تعدیل یافته در ۱۰۰۰ مرتبه تکرار آزمون

FDR	TP	FP	$\alpha$	FDR	TP	FP	روش	حجم نمونه	$\alpha$
۰/۰۴۴	۳۲۷	۱۴	۰/۰۳	۰/۰۶۷	۴۲۹	۳۱	PP	(۶،۶،۶)	۰/۰۵
۰/۰۵۵	۳۶۱	۲۱		۰/۰۹۹	۴۳۴	۴۸	UP		
۰/۰۲۳	۴۳۸	۱۵		۰/۰۶۵	۵۱۱	۳۶	PP	(۱۰،۱۰،۷)	
۰/۰۴۹	۶۴۰	۳۳		۰/۰۹۷	۵۸۶	۶۳	UP		
۰/۰۳۹	۳۴۴	۱۴		۰/۰۶۷	۴۴۵	۳۲	PP	(۱۲،۱۰،۷)	
۰/۰۵۱	۳۷۲	۲۰		۰/۰۹۵	۴۳۸	۴۶	UP		

بر روی داده‌های واقعی که مربوط به بیماری آنزایمر است نیز این دو روش پیاده‌سازی شده و مقدار نرخ کشف کاذب برای دو روش PP و UP در جدول آمده است.

جدول شماره ۲: نرخ کشف کاذب در دو روش جایگشتی مرسوم و روش جایگشتی تعدیل یافته در داده‌های بیماری آنزایمر

FDR	TP	FP	$\alpha$	FDR	TP	FP	روش	حجم نمونه	$\alpha$
۰/۰۴۱	۹۰۹	۳۹	۰/۰۳	۰/۰۷۰	۱۰۳۱	۷۸	PP	(۱۰،۱۰،۹)	۰/۰۵
۰/۰۵۹	۱۰۳۳	۶۵		۰/۰۹۳	۱۱۷۶	۱۲۱	UP		



## بحث

روش‌های آماری بسیاری وجود دارد که می‌توان از آن‌ها در مقایسه گروه کنترل و چندین گروه آزمایشگاهی بهره برد و اختلافات معنی‌دار را تشخیص داد، ولی توجه به این نکته بسیار حایز اهمیت است که به کار بردن روش‌های سنتی در داده‌های بزرگ مقیاس از جمله داده‌های ریز آرایه، باعث می‌گردد تا خطا از کنترل خارج شود و این امر خود منجر به گزارش اختلافات معنی‌داری می‌شود که به‌واقع درست نبوده و لذا مقدار FP افزایش می‌یابد (۱).

به لحاظ آماری، یکی از دلایل مهم برای افزایش خطای تصمیم‌گیری در مورد پذیرش یا عدم پذیرش تأثیر یک ژن در بروز یک بیماری که خود مبنای تحلیل بسیاری از نرم‌افزارها نیز است، دقیق نبودن تخمین مقدار معنی‌داری در روش‌های سنتی و در مقایسه‌هایی با تعداد بالا می‌باشد؛ بنابراین لزوم استفاده و آشنایی با روش‌های به‌روزتر و کارآمدتر برای تحلیل چنین داده‌هایی بسیار حیاتی است (۱۶، ۱۲). در این مقاله سعی گردید تفاوت آزمون فرض جایگشتی مرسوم فیشر با آزمون فرض جایگشتی تعدیل یافته در تعیین نرخ کشف کاذب بررسی گردد؛ چراکه این روش، تخمین دقیق‌تری از مقدار معنی‌داری حاصل می‌کند و به‌صورت مستقیم نیز در کاهش مقدار FP و کنترل نرخ کشف کاذب مؤثر است. در این مقاله، نرخ کشف کاذب در مقایسه گروه کنترل و چندین گروه آزمایشگاهی و با استفاده از دو روش مطرح‌شده در مقاله موردبررسی قرار گرفت. در این راستا، با اختیار سطوح مختلف برای خطای اسمی  $\alpha$  و تعیین نرخ کشف کاذب در هر یک از دو روش مطرح‌شده، نشان داده شد این نرخ در روش مرسوم و برای تعداد مقایسه‌های زیاد که معمولاً در داده‌های بزرگ مقیاس رخ می‌دهد، با ضعف همراه است. Blalock (۱۳) برای مقایسه ۸۷۹۹ ژن موردبررسی در روند افزایش سن مغز و تأثیر آن بر بیماری آلزایمر، روش غربالگری را برمبنای آزمون فیشر مرسوم به کار برد تا تعداد مقایسه‌ها را به ۱۹۸۵ مقایسه کاهش دهد و بتواند نرخ کشف کاذب را کنترل کند. Blalock در سطح خطای اسمی کمتر از مقدار تقریبی ۰/۰۳، تعداد FP را عدد ۲۳۳ گزارش کرد و اذعان داشت برای کنترل نرخ کشف کاذب، انتظار داشته است تا تنها ۵۰ عدد FP

در مجموع ژن‌های گزارش شده، مشاهده کند؛ بنابراین بیان می‌کند فاصله این عدد تا ۲۳۳ زیاد است. Dai و Charnigo (۱۴) با استفاده از غربالگری و آزمون‌های مرسوم آماری، ۲۸۶۰ ژن را در بروز بیماری آلزایمر تشخیص دادند که نشان داد تعداد FP در این روش و در سطح خطای اسمی ۰/۰۵ و ۰/۰۳ به ترتیب ۲۹۷ و ۱۷۸ ژن بوده است. در مقابل، با شبیه‌سازی نیز مشخص گردید روش به کار رفته در این مقاله، این تعداد را در سطوح مختلف نمونه در حدود ۱۵ عدد نشان می‌دهد که به مقدار مورد انتظار بسیار نزدیکتر است، در ضمن مقدار نرخ کشف کاذب را نزدیک به مقدار خطای اسمی  $\alpha$  تخمین می‌زند؛ بنابراین آزمون جایگشتی تعدیل یافته، با تخمین مناسب و دقیق‌تر مقدار معنی‌داری، شرایطی را پدید می‌آورد تا مقدار نرخ کشف کاذب را از روش سنتی و معمول بهتر کنترل کند. از آنچه بحث شد می‌توان نتیجه گرفت هر دو روش در شناسایی و تشخیص اختلافات معنی‌دار از یک مکانیسم، یعنی محاسبه مقدار معنی‌داری استفاده می‌کنند. ولی تفاوت اصلی این دو روش در تخمین مقدار معنی‌داری است که روش جایگشتی تعدیل یافته، این مقدار را بسیار دقیق‌تر از روش مرسوم فیشر تخمین می‌زند؛ لذا به نتایج آزمون‌ها با اطمینان بیشتری می‌توان اعتماد کرد.

## نتیجه‌گیری

نتایج شبیه‌سازی نشان داد روش جدید برای مقایسه گروه کنترل با دو گروه آزمایشگاهی، نرخ کشف کاذب را دقیق‌تر کنترل کرده و این مقدار به مقدار اسمی خطای در نظر گرفته‌شده در آزمون نزدیکتر است. با توجه به نتایج به‌دست آمده در شبیه‌سازی، می‌توان انتظار داشت تعداد FP در داده‌های واقعی مربوط به بیماری آلزایمر که از روش PP به‌دست آمده، کمتر باشد و این روش، نرخ کشف کاذب را در سطح اسمی خطای ۰/۰۵ دقیق‌تر از روش UP کنترل می‌کند.

## تشکر و قدردانی

بدین وسیله نویسندگان بر خود لازم می‌دانند تا از نظرهای سردبیر و داوران مقاله که موجب شد این مقاله بهبود پیدا کند تقدیر و تشکر نمایند.

همچنین از کمک‌های فکری استاد برجسته دانشگاه صنعتی کویزنلد استرالیا، سرکار خانم Kerrie Mengersen و از اطلاعات و کمک‌های فکری Eric Blalock از دانشگاه کنتاکی کمال تشکر را داریم.

## References:

1. Efron B. Empirical Bayes estimates for large-scale prediction problems. *J Am Stat Assoc* 2009;104(487):1015–28. PubMed
2. Pan W. On the use of permutation in and the performance of a class of nonparametric methods to detect differential gene expression. *Bioinformatics* 2003;19(11):1333–40. PubMed
3. Efron B, Tibshirani R, Storey JD, Tushern V. Empirical Bayes analysis of a microarray experiment. *J Am Stat Asso* 2001;96(456):1151–60. Link
4. Storey JD, Tibshirani R. Statistical significance for genome wide studies. *Proc Nati Acad Sci U S A*. 2003;100(16):9440–5. PubMed
5. Xie Y, Pan W, Khodursky AB. A note on using permutation based false discovery rate estimates to compare different analysis methods for microarray data. *Bioinformatics* 2005;21(23):4280–88. PubMed
6. Xu J, Cui X. Robustified MANOVA with applications in detecting differentially expressed genes from oligonucleotide arrays. *Bioinformatics* 2008;24(8):1056–62. PubMed
7. Gao X. Construction of null statistics in permutation-based multiple testing for multi-factorial microarray experiments. *Bioinformatics* 2006;22(12):1486–94. PubMed
8. Zhao J, Pan W. Modified nonparametric approaches to detecting differentially expressed genes in replicated microarray experiments. *Bioinformatics* 2003;19(9):1046–54. PubMed
9. Southworth LK, Kim SK, Owen AB. Properties of balanced permutations. *J Comput Biol* 2009;16(4):625–38. PubMed
10. Pesarin F, Salmaso L, Carozzo E, Arboretti R. Union–intersection permutation solution for two-sample equivalence testing. *Stat Comput* 2016;26(3):693–701. Link
11. Bansal NK, Hamedani GhG, Maadooliat M. Testing multiple hypotheses with skewed alternatives. *Biometrics* 2016;72(2):494–502. Link
12. Salehi M, Mohammadpour A, Mohammadi M, Aminghafari M. A modified two-sample t-test based on permutation method for large-scale data. *Communications in Stat-Simulation Comput* 2019;48(2):372-84. Link
13. Blalock EM, Chen KC, Sharrow K, Herman JP, Porter NM, Foster TC, et al. Gene microarrays in hippocampal aging: statistical profiling identifies novel processes correlated with cognitive impairment. *J Neurosci* 2003;23(9):3807–19. PubMed
14. Dai H, Charnigo R. Omnibus testing and gene filtration in microarray data analysis. *J Appl Stat* 2008;35(1):31-47. Link
15. Zhao S, Fung-Leung W, Bittner A, Ngo K, Liu X. Comparison of RNA-seq and microarray in transcriptome profiling of activated T cells. *PloS One* 2014;9(1):e78644. PubMed
16. Salehi M, Mohammadpour A, Mohammadi M, Aminghafari M. A modified f-test for hypothesis testing in large-scale data. *J Biopharm Stat* 2018;28(6):1078-89. PubMed