آزمون‌های پرتوان‌تر از آزمون نسبت درستنمایی با استفاده از فرضیه‌های اجتماع-اشتراک برای واریانس نمونه‌های مستقل از توزیع نرمال

زهرا نیکنام، رحیم چینی‌پرداز

گروه آمار، دانشگاه شهید چمران اهواز

چکیده: آزمون‌های فرضیه کلاسیک، در حالتی که پارامترهای تحت آزمون دارای محدودیت نباشند، آزمون‌های مناسب، مانند آزمون‌های به طور یکنواخت پرتوان‌ترین و آزمون‌های به طور یکنواخت پرتوان‌ترین نااریب را ارائه می‌دهند. این آزمون‌ها برای فرضیه‌های خاص مانند یک‌طرفه و دوطرفه برای پارامتر شکل گرفته‌اند. امّا در عمل ممکن است با فرضیه‌هایی مواجه شویم که پارامترهای تحت آزمون، دارای نوعی محدودیت در فرضیه صفر و یا در فرضیه مقابل باشند. چنین فرضیه‌هایی در چارچوب آزمون فرضیه‌های کلاسیک نمی‌گنجند. بنابراین آماردانان به جای پرتوان‌ترین آزمون‌ها، به دنبال آزمون‌های پرتوان‌تر هستند. در مقاله حاضر آزمون اجتماع-اشتراک برای آزمون علامت واریانس‌های چند جامعه نرمال به دست آمده و با آزمون نسبت درستنمایی مقایسه شده است. با وجود اینکه آزمون اجتماع-اشتراک پرتوان‌تر است، امّا هر دو آزمون نااریب نیستند. دو آزمون مستطیلی و هموار کننده جهت آزمون با توان بیشتر مورد بررسی قرار گرفته‌اند.

واژه‌های کلیدی: آزمون اجتماع-اشتراک، آزمون نسبت درستنمایی، آزمون پرتوان‌تر، آزمون مستطیلی، آزمون هموار .

آدرس الکترونیکی نویسنده مسئول مقاله: رحیم چینی‌پرداز، chinipardaz_ r@scu.ac.ir

۱ مقدمه

فرض کنید برای $i = ۱, \ldots, p$ $(p \geq ۲)$ نمونه‌های تصادفی مستقل $X_{i۱}, \ldots, X_{in_i}$ $(n_i \geq ۲)$ از توزیع F_{θ_i} باشند. می‌خواهیم فرضیه‌های

$$\begin{cases} H_۰ : \theta_i \leq \theta_{i۰}, & \text{برای بعضی از } \theta_i \text{ ها} \\ H_۱ : \theta_i > \theta_{i۰}, & \text{برای همه } \theta_i \text{ ها} \end{cases}$$

را آزمون کنیم. در اینجا $\theta_{i۰}$ها مقادیر ثابت مثبت و معلوم فرض می‌شوند. برای چنین فرضیه‌هایی، آزمون‌های به طور یکنواخت پرتوان‌ترین (UMP) و آزمون‌های به طور یکنواخت پرتوان‌ترین نااریب $(UMPU)$ موجود نیستند (لیو و سینها ، ۱۹۹۵). متداول‌ترین آزمون‌ها برای این فرضیه‌ها، آزمون‌های نسبت درستنمایی تعمیم‌یافته (GLR) هستند. آزمون‌های GLR برای چنین فرضیه‌هایی در مقایسه میانگین‌های توزیع نرمال یک متغیره و چندمتغیره در ساسابوچی (۱۹۸۰، ۱۹۸۸) و در حالت محدودیت ترتیبی میانگین‌ها در ساسابوچی (۲۰۰۷) به کار گرفته شده‌اند. یک بررسی دقیق این آزمون‌ها را می‌توان در پرلمن و وو (۱۹۹۹) دید. ایراد اساسی به آزمون‌های GLR بهینه نبودن آن‌هاست. مثال‌های ساده‌ای وجود دارد که بیانگر UMP نبودن آزمون‌های GLR است (لهمن و رومونو، ۲۰۰۶). حتی برای فرضیه‌های پیچیده‌تر، این آزمون‌ها منجر به آزمون‌های اریب می‌شوند که در آن‌ها احتمال رد کردن فرضیه صفر در حالت درست بودن، از رد کردن نادرست آن، بزرگتر است (برگر ، ۱۹۸۹). با وجود این، آزمون‌های GLR هنوز به دلیل سادگی و امکان به دست آوردن تقریب مجانبی آماره آزمون، پرکاربرد هستند.

روش دیگر حل چنین مسئله‌ای، استفاده از قوانین اجتماع–اشتراک و اشتراک–اجتماع در تشکیل فرضیه‌ها است. بردار پارامتر $\boldsymbol{\theta} = (\theta_۱, \ldots, \theta_p)'$ با فضای پارامتر Θ در نظر بگیرید. آزمون فرضیه‌ای

کد موضوع‌بندی ریاضی (۲۰۱۰): 62H15، 62F03.

به‌صورت

$$\begin{cases} H_0 : \cup_{i=1}^{p}(\theta_i \in \Theta_{i0}), \\ H_1 : \cap_{i=1}^{p}(\theta_i \in \Theta_{i0}^c), \end{cases} \tag{1}$$

یک آزمون فرضیه اجتماع–اشتراک (IUT) [1] و آزمون فرضیه‌ای به صورت

$$\begin{cases} H_0 : \cap_{i=1}^{p}(\theta_i \in \Theta_{i0}), \\ H_1 : \cup_{i=1}^{p}(\theta_i \in \Theta_{i0}^c), \end{cases} \tag{2}$$

یک آزمون فرضیه اشتراک–اجتماع [2] (UIT) نامیده می‌شود. آزمون‌های اجتماع–اشتراک (IUT) به دلیل کاربرد بیشتر، نسبت به UIT توسعه بیشتری یافته‌اند. یک صورت از آزمون‌های اجتماع–اشتراک با بازتعریف (۱) به‌صورت $H_0 : \min_{1\leq i\leq p}\{\theta_i \leq \theta_{i0}\}$ در مقابل $H_1 : \min_{1\leq i\leq p}\{\theta_i > \theta_{i0}\}$ است که آزمون علامت گفته می‌شود و مورد توجه آماردانان از جمله لهمن (۱۹۵۲)، گاتمن (۱۹۸۷)، شرلی (۱۹۹۲) و لیو و سینها (۱۹۹۵) قرار گرفته است. این آماردانان آزمون علامت را در بعضی از توزیع‌های آماری مورد بررسی و نشان داده‌اند که روش آن‌ها منجر به یافتن آزمون پرتوان‌تر از آزمون GLR می‌شود. گاتمن (۱۹۸۷) برای $p = 2$ در دو توزیع از خانواده مکانی و شرلی (۱۹۹۲) برای $p = 3$ آزمون پرتوان‌تر را پیشنهاد داده‌اند. برگر و سینکلایم (۱۹۸۴) با تقسیم فرضیه‌های آماری به زیر فضاهای جامعه نرمال، پرتوان‌تر بودن IUT را نسبت به آزمون‌های GLR اثبات و در این مقالات، همچنین آزمون‌های GLR مورد نقد قرار گرفته‌اند. سایکلی و برگر (۲۰۰۲) نشان داده که برای آزمون‌های علامت در میانگین‌های چند جامعه نرمال، آزمون GLR توان کمتری نسبت به IUT دارد.

چان و همکاران (۲۰۱۵) آزمون‌های جدیدی برای مقایسه پارامتر مستقل نمونه‌های مستقل از توزیع گاما ارائه دادند که ناحیه رد آن‌ها مشابه لیو و برگر (۱۹۹۵)، برگر (۱۹۹۶)، سایکلی و برگر (۲۰۰۲) است. لیو و ژو (۲۰۱۰) و وو و همکاران (۲۰۱۸) فرضیه‌های اجتماع–اشتراک را برای واریانس

[1] Intersection-Union test
[2] Union-Intersection test

توزیع نرمال بر اساس p–مقدار تعمیم‌یافته[۳] (GPV) مورد بررسی قرار دادند و با استفاده از مطالعات شبیه سازی به مقایسه اندازه و توان آزمون روش GPV و GPV جایگزین با آزمون GLR، آزمون گاتمن (۱۹۸۷)، آزمون لیو و سینها (۱۹۹۵) و آزمون لیو و چان (۲۰۱۰) پرداختند. در تمام موارد، آماردانان ناحیه‌هایی به ناحیه رد آزمون GLR اضافه کرده‌اند به گونه‌ای که تحت فرضیه H_0 تابع توان یا خطای آزمون در همان سطح α ثابت بماند در حالی که به دلیل اضافه شدن ناحیه‌های اضافی به ناحیه رد، توان آزمون افزایش می‌یابد. هرگاه آماره‌های آزمون مستقل از یکدیگر و فرضیه‌های آماری متعامد باشند این ناحیه‌ها به صورت تعدادی مستطیلی خواهند بود که به همین دلیل به روش مستطیلی مرسوم است (برگر ، ۱۹۸۹؛ لیو و سینها ، ۱۹۹۵). امّا اگر فرضیه‌ها غیرمتعامد باشند ممکن است این ناحیه‌های به صورت‌های دیگر مانند لوزی باشند (برگر ، ۱۹۸۹). در روش هموار نیز اضافه کردن مجموعه‌های مجزا به ناحیه رد به صورت اضافه کردن ناحیه‌های اضافی به ناحیه رد آزمون GLR به صورت مجزا انجام می‌شود به صورتی که سطح آزمون در همان مقدار اولیه باقی بماند (لیو و برگر ، ۱۹۹۵؛ مک‌دمارت و وانگ، ۲۰۰۲).

هدف این مقاله، بررسی و مقایسه توان‌های آزمون در واریانس‌های چند جامعه نرمال است که در آن علاوه بر معرفی دو روش آزمون GLR و IUT و مقایسه آن‌ها، دو آزمون پرتوان‌تر نیز معرفی خواهند شد. لازم به ذکر است که چان و همکاران (۲۰۱۵) برای مقایسه پارامترهای دو توزیع گاما روش هموار را مطرح کرده و اشاره کرده‌اند که روش آنان برای مقایسه واریانس‌های جامعه نرمال قابل استفاده است. وو و همکاران (۲۰۱۸) نیز با کمک شبیه سازی و به‌دست آوردن p–مقدار دو واریانس را مقایسه کرده است. در این مقاله علاوه بر روش اجتماع– اشتراک و مقایسه آن با روش مرسوم GLR هر دو روش هموار و مستطیلی برای مقایسه واریانس‌های دو جامعه نرمال در حالت کلی انجام شده و به صورت عددی برای حالت خاص $p = ۲$ چهار روش مقایسه شده‌اند. از آنجا که مقایسه واریانس‌های جامعه‌های نرمال حتی در مقایسه میانگین توزیع نرمال نیز تعیین کننده است، اهمیت این بررسی را بیشتر می‌کند. بدیهی است همانند تحقیقات پیشین اگر آزمونی برای فرضیه‌های (۱) پرتوان‌تر از آزمون‌های GLR و با محاسبات ساده به‌دست آید، مورد توصیه محققین قرار می‌گیرد. در بخش ۲، ارتباط بین آزمون GLR و IUT بیان

[3] Generalized p-value

می‌شود. در بخش ۳، آزمون علامت واریانس‌ها در چند جامعه نرمال به روش GLR و IUT به‌دست آورده شده و با هم مقایسه می‌شوند. در بخش‌های ۴ و ۵، آزمون مستطیلی و آزمون هموار کننده، برای آزمون علامت واریانس‌های دو جامعه نرمال نمونه‌های مستقل به کار می‌رود. مقایسه این دو آزمون نیز در همین بخش آمده است. بخش ۶، به مقایسه توان این آزمون‌ها اختصاص داده شده است. در نهایت بحث و نتیجه‌گیری نتایج بخش‌های قبلی، ارائه می‌شود.

۲ آزمون GLR و IUT

آزمون فرضیه (۱) را در نظر بگیرید. برگر (۱۹۸۹، ۱۹۹۶) با این استدلال منطقی که فرضیه H_0 رد می‌شود اگر همه فرضیه‌های $H_{10}, \ldots, H_{p0}$ رد شوند، ناحیه رد IUT را به‌صورت $R = \cap_{i=1}^{p} R_i$ تعریف کرد، که در آن R_i ناحیه رد آزمون $H_{i0} : \theta_i \leq \theta_{i0}$ در مقابل $H_{i1} : \theta_i > \theta_{i0}$ است. اگرچه در IUT، تاکیدی بر روش به‌دست آوردن ناحیه رد R_i نیست. امّا چون برای فرضیه‌های یک‌طرفه معمولاً و در مطالعه حاضر آزمون GLR به آزمون‌های UMP منجر می‌شوند می‌توان برای آزمون‌های منفرد به روش GLR اکتفا کرد. برگر (۱۹۹۷) اثبات کرده است که اگر همه آزمون‌های منفرد $H_{i0} : \theta_i \leq \theta_{i0}$ در مقابل $H_{i1} : \theta_i > \theta_{i0}$ دارای سطح α باشند، آنگاه آزمون با ناحیه رد R برای H_0 در مقابل H_1 همچنان با سطح α خواهد بود. همچنین، فرض کنید یکی از این آزمون‌های منفرد، مانند i، با سطح α باشد و برای هر آزمون دیگر $j \neq i\,(j = 1, \ldots, p)$ با اندازه α، و ناحیه رد R_j با افزایش θ_j برای H_{j0} در مقابل H_{j1} دنباله $\theta_\ell \in \Theta_i$ در دو رابطه

$$\lim_{\ell \to \infty} P_{\theta_\ell}(\boldsymbol{X} \in R_i) = \alpha, \qquad \lim_{\ell \to \infty} P_{\theta_\ell}(\boldsymbol{X} \in R_j) = 1, \tag{۳}$$

صدق کند، آزمون با ناحیه رد R برای H_0 در مقابل H_1 با اندازه α خواهد بود. این نتایج نشان می‌دهد که می‌توان از ترکیب آزمون‌های منفرد H_{i0} در مقابل H_{i1}، IUT به دست آورد. ضمن اینکه با کنترل خطای نوع اول و اندازه آزمون‌های منفرد سطح و اندازه آزمون اصلی را نیز کنترل کرد.

در یک آزمون GLR آماره آزمون برای فرضیه‌های (۱) به‌صورت

$$\begin{aligned}\lambda(\boldsymbol{x}) &= \frac{\sup_{\theta\in\Theta_0} L(\theta|\boldsymbol{x})}{\sup_{\theta\in\Theta} L(\theta|\boldsymbol{x})} = \frac{\max_{1\le i\le p}\sup_{\theta\in\Theta_i} L(\theta|\boldsymbol{x})}{\sup_{\theta\in\Theta} L(\theta|\boldsymbol{x})}\\ &= \max_{1\le i\le p}\frac{\sup_{\theta\in\Theta_i} L(\theta|\boldsymbol{x})}{\sup_{\theta\in\Theta} L(\theta|\boldsymbol{x})} = \max_{1\le i\le p}\lambda_i(\boldsymbol{x}),\end{aligned}$$

تعریف می‌شود، که در آن $\lambda_i(\boldsymbol{x}) = \frac{\sup_{\theta\in\Theta_i} L(\theta|\boldsymbol{x})}{\sup_{\theta\in\Theta} L(\theta|\boldsymbol{x})}$، آماره آزمون GLR برای آزمون فرضیه : H_{i0} $\theta\in\Theta_i$ در مقابل H_{i1} : $\theta\in\Theta_i^c$ است. باید توجه کرد که اگرچه $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_i} L(\boldsymbol{\theta}|\boldsymbol{x})$ با استفاده از تمامی مشاهدات به دست می‌آید و Θ کل فضای پارامتر است مستقل بودن نمونه‌ها ایجاب می‌کند برای برآورد θ_i فقط نمونه $x_{i1},\ldots,x_{in_i}$ مورد استفاده قرار گیرد. بنابراین آماره آزمون GLR برای آزمون فرضیه H_0 در مقابل H_1 به‌صورت $\lambda(\boldsymbol{x}) = \max_{1\le i\le p}\lambda_i(\boldsymbol{x})$ است. با توجه پیچیدگی آماره آزمون و در نتیجه توزیع آن، به دست آوردن ناحیه رد آزمون GLR مشکل است. امّا برای آزمون‌های منفرد ناحیه رد به سادگی به دست می‌آید. راه حل پیشنهادی استفاده از پیشنهاد (برگر و سینکلایم ، ۱۹۸۴) در استفاده از روش IUT است. فرض کنید ناحیه رد آزمون منفرد با اندازه α از $R_i = \{\boldsymbol{x}; \lambda_i(\boldsymbol{x}) < c_{i\alpha}\}$ با $\sup_{\boldsymbol{\theta}\in\Theta_{i0}} P(\lambda_i(\boldsymbol{x}) < c_{i\alpha}) = \alpha$ باشد. ناحیه رد آزمون به‌صورت $R = \cap_{i=1}^k R_i$ در نظر گرفته شود، با توجه به انتخاب ناحیه رد برای آزمون‌های منفرد با اندازه α سطح آزمون IUT هم α خواهد بود. البته برای اینکه نشان داده شود آزمون با اندازه α است، نشان دادن رابطه (۳) ضروری است. اگر $c_{1\alpha} = \cdots = c_{p\alpha} = c$ باشد، آنگاه ناحیه رد IUT و ناحیه رد آزمون GLR یکسان است. چون

$$\begin{aligned}\text{ناحیه رد } IUT &= \cap_{i=1}^p\{\boldsymbol{x}; \lambda_i(\boldsymbol{x}) < c_{i\alpha}\} = \cap_{i=1}^p\{\boldsymbol{x}; \lambda_i(\boldsymbol{x}) < c\}\\ &= \{\boldsymbol{x}; \max_{1\le i\le p}\lambda_i(\boldsymbol{x}) < c\} = \{\boldsymbol{x}; \lambda(\boldsymbol{x}) < c\} = \text{ناحیه رد } GLR.\end{aligned}$$

۳ آزمون GLR و IUT برای مقایسه واریانس‌های توزیع نرمال

فرض کنید به ازای هر $i = ۱, \ldots, p$ $(p \geq ۲)$، $X_{i۱}, \ldots, X_{in_i}$ متغیرهای تصادفی از توزیع نرمال به‌صورت $X_{ij} \sim N(\mu_i, \sigma_i^۲)$ $(j = ۱, \ldots, n_i$ $(n_i \geq ۲))$ باشد. آزمون علامت را برای فرضیه‌های

$$\begin{cases} H_۰ : \bigcup_{i=۱}^{p} \{\sigma_i^۲ \leq \sigma_{i۰}^۲\}, \\ H_۱ : \bigcap_{i=۱}^{p} \{\sigma_i^۲ > \sigma_{i۰}^۲\}, \end{cases} \tag{۴}$$

در نظر می‌گیریم، که در آن $\sigma_{i۰}^۲$ها مقادیر ثابت مثبت نامعلوم هستند. وقتی که μ_iها معلوم باشند نتایج به‌صورت ساده‌تر و مشابه به دست خواهند آمد. آماره آزمون GLR برای فرضیه $H_{i۰} : \sigma_i^۲ \leq \sigma_{i۰}^۲$ در مقابل $H_{i۱} : \sigma_i^۲ > \sigma_{i۰}^۲$ به صورت

$$\lambda_i(\mathbf{X}) = \frac{\sup_{H_{i۰}} L(\mu_i, \sigma_i^۲|\mathbf{X})}{\sup_{H_{i۰} \cup H_{i۱}} L(\mu_i, \sigma_i^۲|\mathbf{X})},$$

تعریف می‌شود، که در آن $\sup_{H_{i۰} \cup H_{i۱}} L(\mu_۱, \ldots, \mu_p, \sigma_۱^۲, \ldots, \sigma_p^۲|\mathbf{X}) = L(\widehat{\mu}_i, \widehat{\sigma}_i^۲)$ و

$$\sup_{H_{i۰}} L(\mu_i, \sigma_i^۲) = \begin{cases} (۲\pi\sigma_{i۰}^۲)^{\frac{-n_i}{۲}} \exp\{-\frac{\sum_{j=۱}^{n_i}(X_{ij}-\mu_i)^۲}{۲\sigma_{i۰}^۲}\} & \widehat{\sigma}_i^۲ < \sigma_{i۰}^۲, \\ L(\widehat{\mu}_i, \widehat{\sigma}_i^۲) & \widehat{\sigma}_i^۲ > \sigma_{i۰}^۲. \end{cases}$$

بنابراین

$$\lambda_i(\mathbf{X}) \quad = \quad \begin{cases} \left(\frac{\widehat{\sigma}_i^۲}{\sigma_{i۰}^۲}\right)^{\frac{n_i}{۲}} \exp\left\{-\frac{n_i}{۲}\frac{\widehat{\sigma}_i^۲}{\sigma_{i۰}^۲}\right\} & \widehat{\sigma}_i^۲ < \sigma_{i۰}^۲, \\ ۱ & \widehat{\sigma}_i^۲ > \sigma_{i۰}^۲, \end{cases}$$

که در آن $\overline{X}_i = \frac{۱}{n_i}\sum_{j=۱}^{p} X_{ij}$ و $\widehat{\sigma}_i^۲ = \frac{۱}{n_i-۱}\sum_{j=۱}^{n_i}(X_{ij}-\overline{X}_i)^۲$ (روهاتگی و صالح ، ۲۰۰۰). آماره آزمون $W_i = \frac{(n_i-۱)S_i^۲}{\sigma_{i۰}^۲}$ تحت فرضیه $H_۰$ دارای توزیع $\chi_{n_i-۱}^۲$ است. فرض کنید $\chi_{n_i-۱}^۲(\alpha)$ چندک بالایی αام توزیع $\chi_{n_i-۱}^۲$ باشد. فرضیه $H_{i۰}$ در مقابل $H_{i۱}$ رد می‌شود اگر $\lambda_i(\mathbf{x}) \leq \chi_{n_i-۱}^۲(\alpha)$.

به ازای $0 < \alpha < 0/5$، $c_{i\alpha} = \left(\frac{\chi^2_{n_i-1}(\alpha)}{n_i}\right)^{\frac{n_i}{2}} \exp\left\{-\frac{1}{2}\chi^2_{n_i-1}(\alpha)\right\}$، آماره آزمون GLR برای فرضیه H_0 در مقابل H_1 به صورت $\lambda(\mathbf{x}) = \max_{1\leq i\leq p} \lambda_i(\mathbf{x})$ است. فرضیه H_0 در مقابل H_1 رد می‌شود هرگاه $\lambda(\mathbf{x}) \leq c_\alpha$ باشد. نقطه بحرانی به صورت $c_\alpha = \min_{1\leq i\leq p} c_{i\alpha}$ است که با توجه به خاصیت توزیع χ^2 به ازای مقادیر ثابت α، با افزایش n_iها، $c_{i\alpha}$ها افزایش می‌یابند، کوچک‌ترین مقدار c_α متناظر با کمترین مقدار نمونه یعنی $n_{(1)} = \min_{1\leq i\leq p} n_i$ است. بنابراین در آزمون GLR فرضیه H_0 رد می‌شود هرگاه برای $i = 1, \ldots, p$ نامساوی

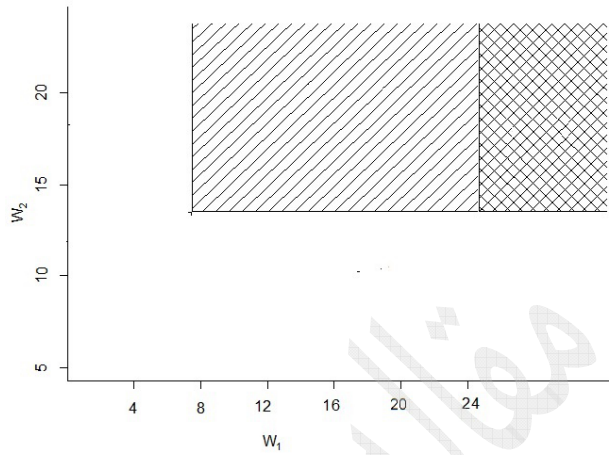
$$W_i \geq n_i\left(\frac{\chi^2_{n_{(1)}-1}(\alpha)}{n_{(1)}}\right)^{\frac{n_{(1)}}{n_i}} \left(\exp\left\{\frac{1}{2}\chi^2_{n_i-1}(\alpha) - \frac{1}{2}\chi^2_{n_{(1)}-1}(\alpha)\right\}\right)^{\frac{2}{n_i}}, \qquad (5)$$

برقرار باشد. یک IUT برای H_0 در مقابل H_1 با اندازه α دارای ناحیه رد

$$W_i \geq \chi^2_{n_i-1}(\alpha), \qquad (6)$$

خواهد بود. اگر $n_1 = \ldots = n_p$ باشد، طرف راست (۵) و (۶) برابر خواهند شد، در نتیجه دو آزمون به یک ناحیه رد منجر خواهند شد. در شکل ۱ ناحیه رد برای دو آزمون GLR و IUT برای $p = 2$، $n_1 = 2$ و $n_2 = 8$ با $\alpha = 0/05$ نشان داده شده است. با توجه به این شکل‌ها ناحیه رد آزمون IUT، $\{W_1 \geq 7/81, W_2 \geq 14/07\}$ و ناحیه رد GLR، $\{W_1 \geq 24/44, W_2 \geq 14/07\}$ است. ناحیه رد آزمون IUT (شامل دو ناحیه هاشور زده) از ناحیه رد GLR بزرگ‌تر است. در اینجا هر دو آزمون GLR و IUT یکنوا[4] هستند چون اگر نقطه‌های نمونه‌ای $w_1, \ldots, w_p$ در ناحیه رد قرار داشته باشند و به‌ازای همه مقادیر $i = 1, \ldots, p$ و $w'_i \geq w_i$، $w'_1, \ldots, w'_p$ نیز در ناحیه رد قرار دارند. همچنین دو آزمون بالا نااریب نیستند. یعنی دارای توان کمتر از اندازه آزمون در ناحیه رد هستند. به عنوان مثال توان آزمون GLR در ناحیه رد و نزدیک به نقطه مرزی دو فرضیه برابر α^2 است. لازم به ذکر است که برای مقادیر دیگر α و همچنین تعداد نمونه‌های دیگر نتیجه کلی، تفاوتی ندارد. همچنین برای تعداد

[4] Monotone



شکل ۱: ناحیه رد آزمون‌های GLR و IUT برای فرضیه‌های آزمون علامت در مقایسه واریانس‌های جامعه نرمال، هاشور یک‌طرفه IUT و هاشور دوطرفه GLR.

نمونه‌های برابر در دو نمونه، ناحیه‌های رد و بنابراین توان دو آزمون یکسان خواهد بود.

۴ آزمون مستطیلی برای واریانس نمونه‌های مستقل در جامعه نرمال

فرض کنید آماره آزمون T_i دارای تابع توزیع تجمعی $F_i(t)$ باشد. اگر $0 < \alpha < \frac{1}{2}$ و عدد صحیح J تعریف شود به طوری که در رابطه $J - 1 < \frac{1}{2\alpha} \leq J$ صدق کند. با تعریف مقادیر ثابت

$$c_j^i = t_i^U = \begin{cases} F_i^{-1}(1 - j\alpha) & 1 \leq J, \\ \infty & j = 0, \\ m_i & j = J, \end{cases}$$

که در آن $m_i; i = 1, \ldots, p$، میانه توزیع F_i است. برای مقادیر معمول $\alpha = 0/01, 0/05, 0/1$ مقدار انتهایی عدد صحیح $J = \frac{1}{2\alpha}$ خواهد بود. سایکلی و برگر (۲۰۰۲) با در نظر گرفتن

$$R_j = \{(t_1, \ldots, t_p) : c_j^1 \leq t_1 < c_{j-1}^1, \ldots, c_j^p \leq t_p < c_{j-1}^p\},$$

ناحیه رد آزمون مستطیلی[5] را به‌صورت $R = \cup_{j=1}^{J} R_j$ تعریف کرده و نشان دادند $\sup_{\theta\in\Theta_0} P(R) \leq \alpha$. بدون از دست دادن کلیت مسئله، $p = 2$ را برای مسئله آزمون فرضیه‌های (۴) در نظر بگیرید. چون $W_i = \frac{(n_i-1)S_i^2}{\sigma_{i0}^2}$ از توزیع $\chi^2_{n_i-1}$ پیروی می‌کند. پس

$$R_j = \{(W_1, W_2) : \chi^2_{n_1-1}(j\alpha) \leq W_1 < \chi^2_{n_1-1}((j-1)\alpha), \chi^2_{n_2-1}(j\alpha) \leq W_2 < \chi^2_{n_2-1}((j-1)\alpha)\}.$$

در اینجا با تعریف $\chi^2_{n_1-1}(0) = \infty$ برای $j = 1$ با اندکی تفاوت

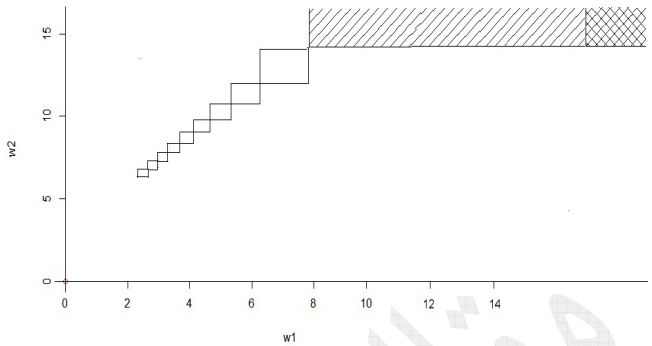
$$R_1 = \{(W_1,\ W_2) : W_1 \geq \chi^2_{n_1-1}(\alpha), W_2 \geq \chi^2_{n_2-1}(\alpha)\},$$

به دست می‌آید. با توجه به تعاریف بالا و همچنین به کارگیری نتایج سایکلی و برگر (۲۰۰۲)، $P_{H_0}((W_1, W_2) \in \cup_{j=1}^{J} R_j) = \alpha$ ناحیه رد آزمون IUT معادل R_1 و در نتیجه توان آزمون مستطیلی با $\alpha = 0/05$ به صورت

$$\begin{aligned}
\beta_R(\sigma_1^2, \sigma_2^2) &= P_{H_1}\big((W_1, W_2) \in (R_1 \cup \cdots \cup R_{10})\big) \\
&= \beta_{IUT} + \sum_{j=2}^{10} P_{H_1}(\mathbf{W} \in R_j) \\
&= \beta_{IUT} + \sum_{j=2}^{10} \big[P_{H_1}\big(\chi^2_{n_1-1}(j\alpha) \leq W_1 < \chi^2_{n_1-1}((j-1)\alpha)\big)\big] \\
&\times \ \big[F_{W_2}\big(\frac{\sigma_{20}^2}{\sigma_2^2}\chi^2_{n_2-1}(j\alpha)\big) - F_{W_2}\big(\frac{\sigma_{20}^2}{\sigma_2^2}\chi^2_{n_2-1}((j-1)\alpha)\big)\big],
\end{aligned}$$

خواهد بود که توان بیشتری نسبت به IUT و بنابراین آزمون GLR دارد. توجه شود در اینجا منظور از $F_{W_i}(u)$ تابع توزیع تجمعی $W_i, i = 1, 2$ در نقطه u یعنی تابع توزیع $\chi^2_{n_i-1}$ در u است. ناحیه رد آزمون GLR، IUT و آزمون مستطیلی برای واریانس‌های نمونه‌های مستقل از توزیع نرمال به ازای $\alpha = 0/05$، $n_1 = 4$ و $n_2 = 8$ در شکل ۲ نشان داده شده است.

[5] Rectangle



شکل ۲: ناحیه رد آزمون‌های GLR، IUT و R برای فرضیه‌های آزمون علامت در مقایسه واریانس‌های جامعه نرمال، هاشور یک‌طرفه IUT و هاشور دوطرفه GLR، مستطیل اول $R_۲$، مستطیل دوم $R_۳$، ... و مستطیل آخر $R_{۱۰}$.

۵ آزمون هموارکننده برای واریانس نمونه‌های مستقل از جامعه نرمال

در این بخش ابتدا آزمون هموارکننده[۶] (S) معرفی می‌شود، سپس ناحیه رد آزمون هموارکننده برای واریانس‌های نمونه‌های مستقل از توزیع نرمال به دست آورده می‌شود. مجموعه A را در مربع واحد $[۰,۱] \times [۰,۱]$ در نظر بگیرید که به سه مجموعه

$$\begin{aligned}
A_۰ &= \{(u_۱,u_۲) : ۱-\alpha \le u_۱ \le ۱,\ ۱-\alpha \le u_۲ \le ۱\},\\
A_۱ &= \{(u_۱,u_۲) : |u_۱-u_۲| \le \frac{\alpha}{۲},\ \frac{۱}{۲} \le u_۱ \le ۱-\alpha,\ \frac{۱}{۲} \le u_۲ \le ۱-\alpha\},\\
A_۲ &= \{(u_۱,u_۲) : \frac{۱}{۲} \le u_۲ \le u_۱ - \frac{۱}{۲} + ۳\frac{\alpha}{۲},\ u_۱ < ۱-\alpha\}\\
&\quad \cup \{(u_۱,u_۲) : \frac{۱}{۲} \le u_۱ \le u_۲ - \frac{۱}{۲} + ۳\frac{\alpha}{۲},\ u_۲ < ۱-\alpha\},
\end{aligned} \tag{۷}$$

افراز شده است. مک‌دمارت و وانگ (۲۰۰۲) نشان دادند که مجموعه A دارای خاصیتی مهم است. فرض کنید $A(u_۲) = \{u_۱ : (u_۱,\ u_۲) \in A\}$، اگر $u_۲ < \frac{۱}{۲}$، آن‌گاه $A(u_۲) = \emptyset$. اگر $\frac{۱}{۲} \le u_۲ \le ۱$،

[6] Smoother

آن‌گاه $A(u_2)$ شامل یک یا دو فاصله با مجموع طول برابر α خواهد بود. اگر U_1 و U_2 متغیرهای تصادفی مستقل و U_1 دارای توزیع یکنواخت $(0,\ 1)$ باشد، با صرف نظر از توزیع U_2 داریم

$$P((U_1,\ U_2) \in A) = \int_0^1 P(U_1 \in A(u_2)) f(u_2) du_2 = \alpha P(\frac{1}{2} \leq U_2 \leq 1) \leq \alpha.$$

از آنجا که مجموعه A نسبت به U_1 و U_2 متقارن است، این برای U_2 دارای توزیع یکنواخت $(0,\ 1)$ نیز برقرار است، یعنی $P((U_1, U_2) \in A) = \alpha P(\frac{1}{2} \leq U_1 \leq 1) \leq \alpha$. اکنون از این نتایج می‌توان برای فرضیه H_0 در مقابل H_1 در (۴) استفاده کرد. فرض کنید $U_1 = F_1(w_1)$ و $U_2 = F_1(w_2)$ و فرضیه H_0 رد شود هرگاه $(U_1,\ U_2) \in A$. اگر A ناحیه رد آزمون S در نظر گرفته شود یادآوری می‌شود $\{(U_1, U_2) \in A_0\} = \{W_1 \geq c_{1,\alpha},\ W_2 \geq c_{2,\alpha}\}$، چون $F_i(W_i)$ دارای توزیع یکنواخت $[0, 1]$ است و $c_{i,\alpha},\ i = 1, 2$ نقطه‌ای است که سطح زیر منحنی بعد از آن برابر α است. بنابراین آزمون با ناحیه $A = A_0 \cup A_1 \cup A_2$ در حالی که دارای اندازه α خواهد بود. ناحیه رد آزمون شامل ناحیه رد آزمون GLR و IUT و نیز دو ناحیه اضافی دیگر A_1 و A_2 است بنابراین از آن‌ها پرتوان‌تر است. با توجه به رابطه (۷) مجموعه های A_0، A_1 و A_2 به‌صورت

$$\begin{aligned}
A_0 &= \{(W_1, W_2) : W_1 \geq \chi^2_{n_1-1}(\alpha),\ W_2 \geq \chi^2_{n_2-1}(\alpha)\},\\
A_1 &= \{(W_1, W_2) : F_2^{-1}(F_1(W_1) - \frac{\alpha}{2}) \leq W_2 \leq F_2^{-1}(F_1(W_1) + \frac{\alpha}{2}),\\
&\quad m_1 \leq W_1 < \chi^2_{n_1-1}(\alpha),\ m_2 \leq W_2 < \chi^2_{n_2-1}(\alpha)\},\\
A_2 &= \{(W_1, W_2) : m_2 \leq W_2 \leq F_2^{-1}(F_1(W_1) - \frac{1}{2} + 3\frac{\alpha}{2}),\ W_1 < \chi^2_{n_1-1}(\alpha)\}\\
&\quad \cup \{(W_1, W_2) : m_1 \leq W_1 \leq F_1^{-1}(F_2(W_2) - \frac{1}{2} + \frac{3\alpha}{2}),\ W_2 < \chi^2_{n_2-1}(\alpha)\},
\end{aligned}$$

خواهند بود، که در آن $m_i, i = 1, 2$ میانه توزیع F_i، است. همچنین F_i تابع توزیع تجمعی χ^2 با $n_i - 1$ درجه آزادی و $F_i^{-1}(\nu)$ چندک ν ام توزیع $\chi^2_{n_i-1}$ است.

با درگرفتن $U_i = F_i(W_i)$، برای A_0 داریم

$$\begin{aligned} 1-\alpha \le U_1 < 1 \quad &\Rightarrow \quad F_i^{-1}(1-\alpha) \le W_i < F_i^{-1}(1) \\ &\Rightarrow \quad c_{1\alpha} \le W_1 < +\infty \Rightarrow \chi^2_{n_1-1}(\alpha) \le W_1 < +\infty \\ &\Rightarrow \quad W_1 \ge \chi^2_{n_1-1}(\alpha). \end{aligned}$$

برای A_1 داریم

$$\begin{aligned} \mid U_1 - U_2 \mid \le \frac{\alpha}{2} \quad &\Rightarrow \quad \frac{-\alpha}{2} \le U_1 - U_2 \le \frac{\alpha}{2} \Rightarrow \frac{-\alpha}{2} \le F_1(W_1) - F_2(W_2) \le \frac{\alpha}{2} \\ &\Rightarrow \quad F_1(W_1) - \frac{\alpha}{2} \le F_2(W_2) < F_1(W_1) + \frac{\alpha}{2} \\ &\Rightarrow \quad F_2^{-1}(F_1(W_1) + \frac{\alpha}{2}) \le W_2 \le F_2^{-1}(F_1(W_1) + \frac{\alpha}{2}). \end{aligned}$$

همین مسئله را می‌توان برای $F_1^{-1}(F_2(W_2) + \frac{\alpha}{2}) \le W_1 \le F_1^{-1}(F_2(W_2) + \frac{\alpha}{2})$ در نظر گرفت.
از طرفی

$$\frac{1}{2} \le U_1 \le 1-\alpha \Rightarrow \frac{1}{2} \le F_1(W_1) \le 1-\alpha \Rightarrow F_1^{-1}(\frac{1}{2}) \le W_1 \le F_1^{-1}(1-\alpha).$$

با همین استدلال $m_2 \le W_2 \le c_{2\alpha} = \chi^2_{n_2-1}(\alpha)$ و برای A_2 داریم

$$\begin{aligned} \frac{1}{2} \le U_2 \le U_1 - \frac{1}{2} + \frac{3\alpha}{2} \quad &\Rightarrow \quad \frac{1}{2} \le F_2(W_2) \le F_1(W_1) - \frac{1}{2} + \frac{3\alpha}{2} \\ &\Rightarrow \quad F_2^{-1}(\frac{1}{2}) \le W_2 \le F_2^{-1}(F_1(W_1) - \frac{1}{2} + \frac{3\alpha}{2}) \\ &\Rightarrow \quad m_2 \le W_2 \le F_2^{-1}(F_1(W_1) - \frac{1}{2} + \frac{3\alpha}{2}), \end{aligned}$$

$$U_1 < 1-\alpha \Rightarrow F_1(W_1) < 1-\alpha \Rightarrow W_1 < F_1^{-1}(1-\alpha) = c_{1\alpha} = \chi^2_{n_1-1}(\alpha).$$

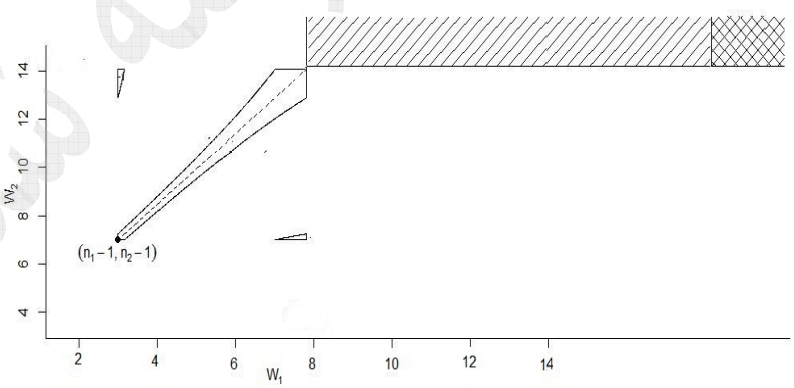
به صورت مشابه برای $U_۲$ داریم

$$U_۲ < ۱ - \alpha \Rightarrow F_۲(W_۲) < ۱ - \alpha \Rightarrow W_۲ < F_۲^{-۱}(۱ - \alpha) = c_{۲\alpha} = \chi^۲_{n_۲ - ۱}(\alpha).$$

در این آزمون، فرضیه $H_۰$ رد می‌شود هرگاه $(W_۱, W_۲) \in A$، یا به بیانی ساده‌تر فرضیه $H_۰$ رد می‌شود هرگاه $U_i = F_i(w_i)$ و $(U_۱, U_۲) \in A$. توان آزمون S برای مدل نرمال به‌صورت

$$\begin{aligned}\beta_S(\sigma^۲_۱, \sigma^۲_۲) &= P_{H_۱}\big((W_۱, W_۲) \in A\big)\\ &= \big(\alpha P_{\sigma^۲_۱, \sigma^۲_{۲۰}}(W_۱ > m_۱)\big)\big(\alpha P_{\sigma^۲_{۱۰}, \sigma^۲_۲}(W_۲ > m_۲)\big)\\ &= \alpha^۲\big(۱ - F_{W_۱}(\frac{m_۱ \sigma^۲_{۱۰}}{\sigma^۲_۱})\big)\big(۱ - F_{W_۲}(\frac{m_۲ \sigma^۲_{۲۰}}{\sigma^۲_۲})\big),\end{aligned}$$

است، که تحت $H_۱$، $\sigma^۲_۲ > \sigma^۲_{۲۰}$ و $\sigma^۲_۱ > \sigma^۲_{۱۰}$ و $m_i,\ i = ۱, ۲$ میانه $F_i(w_i)$ است. به ازای $\alpha = ۰/۰۵$، $n_۱ = ۲$ و $n_۲ = ۸$ ناحیه رد آزمون GLR، IUT و S در شکل ۳ نشان داده شده است.



شکل ۳: ناحیه رد آزمون‌های GLR، IUT و S برای فرضیه‌های آزمون علامت در مقایسه واریانس‌های جامعه نرمال، هاشور یک‌طرفه نشان‌دهنده IUT و هاشور دوطرفه نشان‌دهنده GLR (دو ناحیه هاشور زده $A_۰$) و مخروط شکل $A_۱$ و دو مثلث کوچک $A_۲$ است.

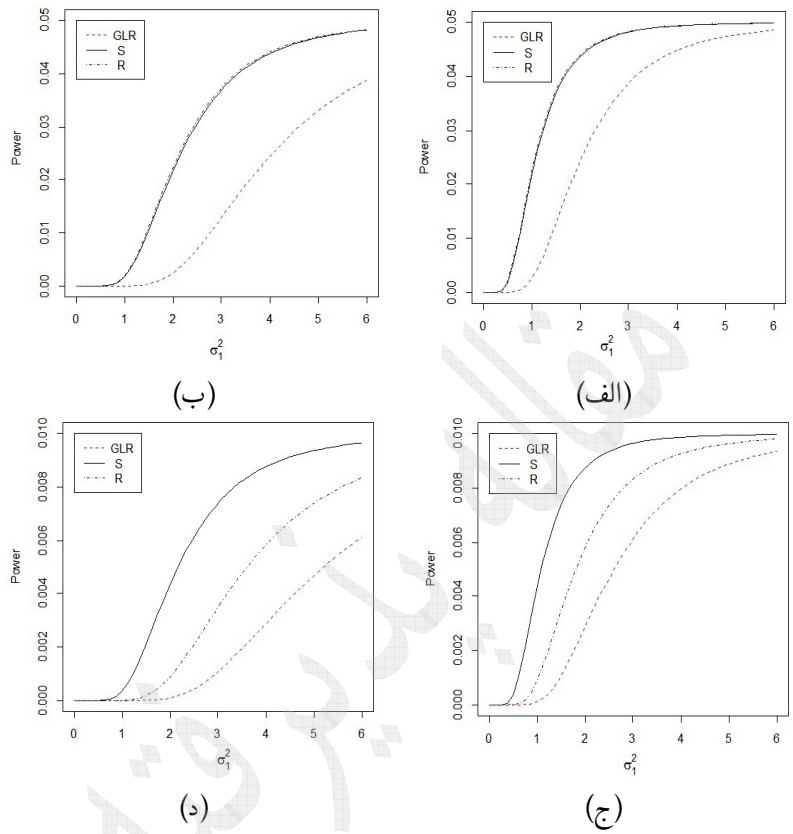
۶ مقایسه توان آزمون‌ها

در این بخش به مقایسه توان‌های سه آزمون اجتماع–اشتراک، مستطیلی، هموارکننده پرداخته می‌شود. در بخش دوم این مقاله نشان داده شد که آزمون GLR توان کمتر و یا مساوی با آزمون IUT دارد. در اینجا نیز بدون از دست دادن کلیت مسئله، فرض کنید $p = ۲$ و $\beta_{IUT}(\sigma_۱^۲, \sigma_۲^۲)$، $\beta_R(\sigma_۱^۲, \sigma_۲^۲)$ و $\beta_S(\sigma_۱^۲, \sigma_۲^۲)$ به ترتیب توان‌های آزمون اجتماع–اشتراک، آزمون مستطیلی و آزمون هموارکننده باشد. توان سه آزمون به ازای مقادیر مختلف پارامترها در شکل ۴ رسم شده است. نتایج عددی این سه آزمون به ازای $n_۱ = ۲$، $n_۲ = ۸$ و مقادیر مختلف پارامترهای $\sigma_۱^۲$، $\sigma_۲^۲$، $\sigma_{۱۰}^۲$، $\sigma_{۲۰}^۲$ و $\sigma_۲^۲(\sigma_۲^۲ = ۰/۵\sigma_۱^۲, \sigma_۱^۲, ۱/۵\sigma_۱^۲)$ در جدول ۱ آورده شده‌اند. نتایج برای مقادیر دیگر که در آن‌ها $\sigma_{۱۰}^۲ > \sigma_{۲۰}^۲$ و یا $\sigma_{۱۰}^۲ < \sigma_{۲۰}^۲$ و حتی مساوی بودن آن‌ها با نمونه‌های مختلف محاسبه شدند که به دلیل مشابه بودن نتایج آورده نمی‌شوند.

جدول ۱: توان IUT، S و R در آزمون علامت دو متغیره با $\alpha = ۰/۰۵$، $n_۱ = ۲$، $n_۲ = ۸$

			$\sigma_۲^۲$					
	آزمون		۳	۲	۱/۷۵	۱/۵	۱/۲۵	۱
	IUT	$(۰/۵\sigma_۲^۲, \sigma_۲^۲)$	۰/۰۷۶۴	۰/۰۲۱۳	۰/۰۱۱۹	۰/۰۰۵۴	۰/۰۰۱۷	۰/۰۰۰۳
	S		۰/۱۴۵۳	۰/۰۹۹۰	۰/۰۸۳۰	۰/۰۶۴۹	۰/۰۴۵۱	۰/۰۲۵۲
	R		۰/۰۹۰۹	۰/۰۴۲۳	۰/۰۳۳۵	۰/۰۲۶۴	۰/۰۲۰۳	۰/۰۱۴۳
$\sigma_{۱۰}^۲ = \sigma_{۲۰}^۲ = ۱$	IUT	$(\sigma_۲^۲, \sigma_۲^۲)$	۰/۱۷۹۹	۰/۰۷۰۵	۰/۰۴۵۶	۰/۰۲۴۸	۰/۰۱۰۲	۰/۰۰۲۵
	S		۰/۲۸۱۷	۰/۲۱۳۱	۰/۱۸۶۶	۰/۱۵۴۴	۰/۱۱۵۹	۰/۰۷۲۴
	R		۰/۱۹۵۱	۰/۰۹۵۶	۰/۰۷۲۶	۰/۰۵۲۳	۰/۰۳۵۵۷	۰/۰۲۲۵
	IUT	$(۱/۵\sigma_۲^۲, \sigma_۲^۲)$	۰/۲۴۸۱	۰/۱۰۹۷	۰/۰۷۴۵	۰/۰۴۳۴	۰/۰۱۹۵	۰/۰۰۵۵
	S		۰/۳۶۳۲	۰/۲۸۵۷	۰/۲۵۴۱	۰/۲۱۴۷	۰/۱۶۵۷	۰/۱۰۷۸
	R		۰/۲۶۲۳	۰/۱۳۴۴	۰/۱۰۱۶	۰/۰۷۱۵	۰/۰۴۶۲	۰/۰۲۶۹
	IUT	$(۰/۵\sigma_۲^۲, \sigma_۲^۲)$	۰/۰۱۶۵	۰/۰۰۲۴	۰/۰۰۱۰	۰/۰۰۰۳	۰/۰۰۰۱	۰/۰۰۰۰
	S		۰/۰۴۵۷	۰/۰۲۵۸	۰/۰۲۰۰	۰/۰۱۴۱	۰/۰۰۸۵	۰/۰۰۳۸
	R		۰/۰۲۶۵	۰/۰۱۴۵	۰/۰۱۲۹	۰/۰۱۱۴	۰/۰۰۹۴	۰/۰۰۶۵
$\sigma_{۱۰}^۲ = ۲, \sigma_{۲۰}^۲ = ۱$	IUT	$(\sigma_۲^۲, \sigma_۲^۲)$	۰/۰۷۶۴	۰/۰۲۱۳	۰/۰۱۱۹	۰/۰۰۵۴	۰/۰۰۱۷	۰/۰۰۰۳
	S		۰/۱۴۵۳	۰/۰۹۹۰	۰/۰۸۳۰	۰/۰۶۴۹	۰/۰۴۵۱	۰/۰۲۵۲
	R		۰/۰۹۰۹	۰/۰۴۲۳	۰/۰۳۳۵	۰/۰۲۶۴	۰/۰۲۰۳	۰/۰۱۴۳
	IUT	$(۱/۵\sigma_۲^۲, \sigma_۲^۲)$	۰/۱۳۳۵	۰/۰۴۶۶	۰/۰۲۸۷	۰/۰۱۴۶	۰/۰۰۵۵	۰/۰۰۱۲
	S		۰/۲۲۳۲	۰/۱۶۲۸	۰/۱۴۰۴	۰/۱۱۴۰	۰/۰۸۳۴	۰/۰۵۰۱
	R		۰/۱۴۸۸	۰/۰۷۰۸	۰/۰۵۴۳	۰/۰۴۰۳	۰/۰۲۸۸	۰/۰۱۹۲

آزمونی که برای مقادیر مختلف $\sigma_۱^۲(\sigma_۱^۲ = ۰/۵\sigma_۲^۲, \sigma_۱^۲ = \sigma_۲^۲, \sigma_۱^۲ = ۱/۵\sigma_۲^۲)$ توانی حداقل برابر با α داشته باشد، آزمونی نااریب در سطح α است. پس آزمون‌های R، S و IUT هر سه اریب هستند.



شکل ۴: توان آزمون‌های GLR، S و R برای فرضیه‌های آزمون علامت در مقایسه واریانس‌های جامعه نرمال الف- به ازای مقادیر $\sigma_{10}^2 = \sigma_{10}^2 = 1, \sigma_2^2 = 1, n_1 = n_2 = 10, \alpha = 0/05$، ب- به ازای مقادیر $\sigma_{10}^2 = \sigma_{10}^2 = 1, \sigma_2^2 = 1, n_1 = 6, n_2 = 10, \alpha = 0/05$، ج- به ازای مقادیر $\sigma_{10}^2 = 2, \sigma_{10}^2 = 1, \sigma_2^2 = 1, n_1 = n_2 = 10, \alpha = 0/05$، د- به ازای مقادیر $\sigma_{10}^2 = 2, \sigma_{10}^2 = 1, \sigma_2^2 = 1, n_1 = 6, n_2 = 10, \alpha = 0/05$

امّا اریبی آزمون R و S از IUT کمتر است. همان طور که انتظار هم می‌رود با افزایش مقدار پارامترها، توان آزمون‌ها افزایش می‌یابد که این اعتماد به نتایج را بیشتر می‌کند. این افزایش در R و S به مراتب از IUT بیشتر است.

بحث و نتیجه‌گیری

اگرچه آزمون‌های GLR در بین روش‌های آزمون کردن متداول‌ترین هستند، امّا برای فرضیه‌های پیچیده، از جمله آزمون‌های علامت که در این مقاله مورد بحث قرار گرفته‌اند، بهینه نیستند. آزمون‌های IUT برای تعداد نمونه نابرابر توان بیشتر و با تعداد نمونه برابر توان مساوی با آن را دارند. بنابراین در مقایسه بین این دو آزمون، IUT با توان بیشتر و محاسبات ساده توصیه می‌شود. هر دو آزمون GLR و IUT توان کمتر از آزمون‌های R و S دارند. بدیهی است اگر امکان استفاده از دو آزمون R و S مقدور باشد بر آزمون‌های قبلی ترجیح داده شوند. در مقایسه دو آزمون R و S نیز نتایج این تحقیق برتری آزمون S بر آزمون R است. چنین نتیجه‌گیری با نتایج آزمون فرضیه علامت در میانگین توزیع‌های نرمال توسط سایکلی و برگر (۲۰۰۲) مطابقت دارد. امّا نتیجه آن در توزیع‌های دیگر و پارامترهای دیگر به نتایج تحقیقات بعدی بستگی دارد. یک ضعف عمده در دو آزمون R و S را می‌توان به ناحیه ردهای مجزا و یکنوا نبودن آن‌ها دانست. چون در آزمون غیر یکنوا، ناحیه‌های رد خاصیت آشیانه‌ای ندارند و نمی‌توان تمام نقاط دورتر از نقاط بحرانی را متعلق به ناحیه رد دانست در چنین شرایطی پیدا کردن ناحیه رد به سادگی صورت نمی‌گیرد و محاسبات را پیچیده می‌کند. به نظر می‌رسد تا زمانی که تحقیقات به محاسبات ساده در این دو آزمون نرسیده است، استفاده از IUT مطلوب‌تر باشد. رسیدن به آزمون‌های UMP و یا حداقل نااریب برای چنین فرضیه‌هایی می‌تواند نقطه عطفی در آزمون‌های آماری باشد.

تقدیر و تشکر

نویسندگان مقاله از سردبیر محترم مجله علوم آماری، هیئت تحریریه، ویراستار محترم و همچنین از داوران ارجمند که با نظرات و پیشنهادهای ارزشمندشان باعث بهبود و اصلاح ارائه این مقاله گردید، کمال تشکر و قدردانی را دارند.

مراجع

Berger, R. L. (1989), Uniformly More Powerful Tests for Hypotheses Concerning Linear Inequalities and Normal Means, *Journal of the American Statistical Association*, **84**, 192-199.

Berger, R. L. (1997), Likelihood Ratio Tests and Intersection-Union Tests, *In Advances in Statistical Decision Theory and Applications*, 225-237. Birkhäuser Boston.

Berger, R. L. and Hsu, J. C. (1996), Bioequivalence Trials, Intersection-Union Tests and Equivalence Confidence Sets, *Statistical Science*, **11**(4), 283-319.

Berger, R. L. and Sinclairm, D. (1984), Testing Hypotheses Concerning Unions of Linear Subspaces, *Journal of American Statistical Association*, **79**, 158-163.

Liu, H. and Chan, C. H. (2010), *More Powerful Tests of the Sign Testing for Normal Variance*, Tech. Rep. No.2010-01, Department of Statistics, National Cheng Chi University, Taipei, Taiwan.

Chan, C. H., Liu, H. and Zen, M. M. (2015), More Powerful Tests for the Sign Testing Problem about Gamma Scale Parameters, *Statistics*, **49**(3), 564-577.

Gail, M. and Simon, R. (1985), Testing for Qualitative Interactions Between Treatment Effects and Patient Subsets, *Biometrics*, **41**, 361- 372.

Gutmann, S. (1987), Tests Uniformly More Powerful than Uniformly Most Powerful Monotone Tests, *Journal of Statistical Planning and Inference*, **17**, 279-292.

Gutmann, S. and Maymin, Z. (1987), Is the Selected Population the Best? *The Annals of Statistics*, **15**, 456-461.

Lehmann, E. L. (1952), Testing Multiparameter Hypotheses, *The Annals of Mathematical Statistics*, **23**, 541-552.

Lehmann, E. L. and Romono, J. P. (2006), *Testing Statistical Hypotheses*, 2ed, Springer, New York.

Liu, H. (1989), Linear Inequality Hypothesis and Uniformly More Power Test, *Journal of China Statistical Association*, **37**, 307-331.

Liu, H. and Berger, R. L. (1995), Uniformly More Powerful, One-Sided Tests for Hypotheses about Linear Inequalities, *The Annals of Statistics*, **23**, 55-72.

Liu, X. and Xu, X. (2010), A New Generalized P-Value Approach for Testing the Homogeneity of Variances, *Staistics and Probability Letters*, **80**, 1486-1491.

Li, T. and Sinha, B. K. (1995), Tests of Ordered Hypotheses for Gamma Scale Parameters, *Journal of Statistical Planning and Inference*, **45**(3), 387-397.

McDermott, M. P. and Wang, Y. (2002), Construction of Uniformly More Powerful Tests for Hypotheses about Linear Inequalities, *Journal of Statistical Planning and Inference*, **107**(1-2), 207-217.

Perlman, M. and Wu, L. (1999), The Emeror's New Tests, with Discussions, *Statistical Science*, **14**(4), 355-381.

Rohatgi, V. K. and Saleh, A. M. E. (2000), *An Introduction to Probability and Statistics*, John Wiley and Sons, New York.

Saikali, K. G. (1996), *Uniformly More Powerful Tests for Linear Inequalities*, Ph.D. Thesis, North Carolina State University, Statistics Department.

Saikali, K. G. and Berger, R. L. (2002), More Powerful Tests for The Sign Testing Problem, *Journal of Statistical Planning and Inference*, **107**(1-2), 187-205.

Sasabuchi, S. (1980), A Test of a Multivariate Normal Mean with Composite Hypotheses Determined by Linear Inequalities, *Biometrika*, **67**(2), 429-439.

Sasabuchi, S. (1988), A Multivariate One-Sided Test with Composite Hypotheses When the Covariance Matrix is Completely Unknown, *Memoirs of the Faculty of Science, Kyushu University. Series A, Mathematics*, **42**, 37-46.

Sasabuchi, S. (1988), More Powerful Tests for Homogenity of Multivariate Normal Mean Vectors under Order Restriction, *Sankhya*, **69**, 700-716.

Shirley, A. G. (1992), Is the Minimum of Several Location Parameters Positive? *Journal of Statistical Planning and Inference*, **31**, 67-79.

Wu, W. Y., Wu, W. H., Hsieh, H. N. and Lee, M. C. (2018), The Generalized Inference on the Sign Testing Problem About the Normal Variances, *Journal of Applied Statistics*, **45**(5), 956-970.

The Uniformly More Powerful Tests than the Likelihood Ratio Test Using Intersection-Union Hypotheses for Variance of Independent Samples from Normal Distribution

Niknam, Z., Chinipardaz, R.

Department of statistics, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

Abstract: Classical hypothesis tests for the parameters provide suitable tests when the hypotheses are not restricted. The best are the uniformly most powerful test and the uniformly most powerful unbiased test. These tests are designed for specific hypotheses, such as one-sided and two-sided for the parameter. However, in practice, we may encounter hypotheses that the parameters under test have typical restrictions in the null or alternative hypothesis. Such hypotheses are not included in the framework of classical hypothesis testing. Therefore, statisticians are looking for more powerful tests than the most powerful ones. In this article, the union-intersection test for the sign test of variances in several normal distributions is proposed and compared with the likelihood ratio test. Although the union-intersection test is more powerful, neither test is unbiased. Two rectangular and smoothed tests have been examined for a more powerful test.

Keywords: Intersection-union test, Likelihood ratio test, More powerful test, Rectangle test, Smoother test

Mathematics Subject Classification (2010): 62F03, 62H15.