مقایسه میانگینهای چند جامعه نرمال با واریانسهای نابرابر

علیرضا زاهدیان[۱] عباس گرامی[۲]

چکیده

آزمون F که برای مقایسه میانگینهای جوامع نرمال به کار می‌رود، به شدت به تساوی واریانسها وابسته است. به عبارت دیگر این آزمون تحت عدول از فرض تساوی واریانسها استوار[۳] نیست. در این مقاله ضمن پرداختن به مسأله بالا (مسألهٔ بهرنس ـ فیشر)، راه‌حلهای پیشنهاد شده برای مقایسه میانگینهای چند جامعهٔ نرمال با واریانسهای نامعلوم و نابرابر مورد مقایسه قرار گرفته و از طریق روش مونت‌کارلو با هم مقایسه شده‌اند.

واژه‌های کلیدی: مسألهٔ بهرنس ـ فیشر، ناهمگنی واریانسها.

۱. مقدمه

رایج‌ترین روش برای انجام آزمون تساوی میانگینهای چند جامعه نرمال، آزمون F (فیشر) می‌باشد [٤]. مانند هر آزمون دیگر، برای بکارگیری این آزمون شرایط خاصی لازم است که یکی از این شرایط، تساوی واریانسهای جوامع مورد بررسی می‌باشد. بدیهی است در صورت فقدان صحت این شرط، آزمون F دچار اشکال می‌شود. از اینرو، آزمونهای پارامتری و ناپارامتری متعددی برای رفع این مشکل پیشنهاد شده‌اند.

در این مقاله علاوه بر آزمون F ، چهار آزمون پیشنهاد شده برای مقایسه میانگینهای چند جامعه نرمال با واریانسهای نابرابر (و نامعلوم) مورد بررسی قرار گرفته، از طریق روش مونت‌کارلو با هم مقایسه شده‌اند و برتری آنها نسبت به آزمون F و هرکدام از نظر اندازه و توان، مشخص شده است.

آزمونهایی که مورد بررسی قرار گرفته‌اند همگی پارامتری هستند و به آزمونهای ناپارامتری پرداخته نشده است زیرا در مقایسه‌ای که توسط بونینگ (۱۹۹۵) انجام شده است، در حالتی که شرط تساوی واریانسها برقرار نیست؛ آزمون ولچ (قسمت (٤.۲)) به عنوان یک آزمون پارامتری نسبت به آزمونهای ناپارامتری، برتر شناخته شده است[۲].

۲. مقایسه میانگینهای چند جامعهٔ نرمال

در حالتی که مقایسهٔ میانگینهای K جامعهٔ نرمال مدنظر باشد، این مقایسه را می‌توان در قالب آزمون فرض $H_0 : \mu_1 = \mu_2 = \cdots = \mu_K$ در مقابل (به ازای حداقل یک $(i \neq j$ ؛ $H_1 : \mu_i \neq \mu_j$، بیان نمود. در شرایطی که واریانسهای

---

[۱] مرکز آمار ایران

[۲] پژوهشکده آمار ایران

[۳] Robust

این K جامعه با هم مساوی باشند، معمول‌ترین روشی که برای آزمون فرض بالا مورد استفاده قرار می‌گیرد، تجزیهٔ واریانس با همان آزمون F است.

۲-۱ آزمون F (تجزیه واریانس)

فرض کنید که مقایسه میانگینهای K جامعه نرمال با وایانسهای برابر مدنظر بوده و از این جامعه‌ها نمونه‌هایی مستقل به حجم $n_1, n_2, \ldots, n_K$ باشرط زیر انتخاب شده باشند

$$X_{ij} \sim N(\mu_i, \sigma^2) \qquad i = 1,2,\ldots,K\,, j = 1,2,\ldots,n_i$$

اگر مجموع مربعات کل، بین نمونه‌ها و داخل نمونه‌ها را به ترتیب زیر تعریف کنیم.

$$SST = \sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}\right)^2 = \sum_{i=1}^{K}\sum_{j=1}^{n_i} X_{ij}^2 - N\overline{X}^2$$

$$SSB = \sum_{i=1}^{K} n_i\left(X_i - \overline{X}\right)^2 = \sum_{i=1}^{K} n_i \overline{X}_i^2 - N\overline{X}^2$$

$$SSE = \sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_i\right)^2 = \sum_{i=1}^{K} n_i S_i^2$$

که

$$\overline{X} = \sum_i \sum_j X_{ij} / N \qquad , \qquad N = \sum_{i=1}^{K} n_i$$

$$S_i^2 = \frac{1}{n_i}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2 \quad و \quad \overline{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$$

تحت فرض H_0 داریم

$$F = \frac{MSB}{MSE} \sim F_{(K-1,N-K)}$$

که $MSB = SSB/(K-1)$ و $MSE=SSE/(N-K)$. بنابراین فرض H_0 در مقابل H_1 در سطح معنی‌داری α رد می‌شود، اگر $F > F_{(K-1,N-K,1-\alpha)}$. در قسمت (۱۰۶۰۲) نشان خواهیم داد که آزمون F نسبت به شرط تساوی واریانسها استوار نیست.

۲-۲ روش ویراهندی (۱۹۹۵)

فرض کنید که مقایسه میانگینهای K جامعه نرمال با واریانسهای نامساوی درنظر باشد. در این صورت نمونه‌های بدست آمده از هر جامعه دارای خواص زیر خواهند بود.

$$X_{ij} \sim N(\mu_i, \sigma_i^2) \quad i = 1,2,\ldots,K \quad j = 1,2,\ldots,n_i$$

فرض کنید که مجموعه‌های مربعات وزنی $\widetilde{SST}$، $\widetilde{SSE}$ و $\widetilde{SSB}$ بصورت زیر تعریف شوند

$$\widetilde{SST} = \sum_{i=1}^{K}\sum_{j=1}^{n_i}\frac{1}{\sigma_i^2}(X_{ij} - \overline{\overline{X}})^2 \tag{1}$$

$$\widetilde{SSE} = \sum_{i=1}^{K}\sum_{j=1}^{n_i}\frac{1}{\sigma_i^2}(X_{ij} - \overline{X})^2 = \sum_{i=1}^{K}\frac{n_i S_i^2}{\sigma_i^2} \tag{2}$$

$$\widetilde{SSB} = \sum_{i=1}^{K} n_i \frac{1}{\sigma^2}(\overline{X}_i - \overline{\overline{X}}_i)^2 = \sum_{i=1}^{K}\frac{n_i \overline{X}_i^2}{\sigma_i^2} - \frac{1}{\Sigma_i n_i / \sigma_i^2}\left(\sum_i \frac{n_i \overline{X}_i}{\sigma_i^2}\right)^2 \tag{3}$$

که $S_i^2 = \frac{1}{n_i}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2$ بوده و $\overline{\overline{X}}$ عبارتست از

$$\overline{\overline{X}} = \frac{\Sigma_i \dfrac{n_i \overline{X}_i}{\sigma_i^2}}{\Sigma_i \dfrac{n_i}{\sigma_i^2}}$$

برای متغیر تصادفی B_j توزیع بتا بصورت زیر برقرار است

$$B_j = \frac{\Sigma_{i=1}^{j}\dfrac{n_i s_i^2}{\sigma_i^2}}{\Sigma_{i=1}^{j+1}\dfrac{n_i s_i^2}{\sigma_i^2}} \sim Beta\left(\frac{\sum_{i=1}^{j}(n_i - 1)}{2}, \frac{n_{j+1} - 1}{2}\right)$$

در رابطه (۳) مجموع مربعات $\widetilde{SSB}$ با وزنهای n_i / σ_i^2 تعریف شد. حال اگر این مجموع مربعات با علامت $\widetilde{SSB} = \widetilde{SSB}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2)$ نشان داده شود، برای وزنهای $n_i \sigma_i^2 s_i^2 / S_i^2$, $(i = 1,2,\ldots,K)$ مجموع مربعات وزنی بین نمونه‌ها

با علامت $S\tilde{S}B\left(s_1^2\sigma_1^2/S_1^2,\ldots,s_K^2\sigma_K^2/S_K^2\right)$ نشان داده می‌شود که در آن s_i^2 برآورد بدست آمده برای S_i^2 براساس مشاهدات نمونه است.

ویراهندی[4] (۱۹۹۵)، $P-value$ تعمیم‌یافته آزمون فرض H_0 در مقابل H_1 را بصورت زیر تعریف کرده است[۶].

$$P-value=\Pr\left(\frac{S\tilde{S}B/(K-1)}{S\tilde{S}E/(N-K)}\ge\frac{N-K}{K-1}\times s\tilde{s}b\left(\frac{n_1s_1^2}{B_1B_2\ldots B_{K-1}},\ldots,\frac{n_Ks_K^2}{1-B_{K-1}}\right)\right)$$
$$=1-H_{(K-1,N-K)}(A)$$

که در آن مقدار بدست آمده برای $S\tilde{S}B$ بر اساس مشاهدات نمونه بوده و A عبارتست از

$$A=\frac{N-K}{K-1}s\tilde{s}b\left(\frac{n_1s_1^2}{B_1B_2\ldots B_{K-1}},\frac{n_2s_2^2}{(1-B_1)B_2\ldots B_{K-1}},\ldots,\frac{n_Ks_K^2}{1-B_{K-1}}\right)$$

و $H_{(K-1,N-K)}(\circ)$ تابع توزیع F با درجات آزادی $K-1$ و $N-K$ می‌باشد.

از آنجایی که در رابطه بالا $1-H_{(K-1,N-K)}(A)$ تابعی از متغیرهای تصادفی B_j ، $(j=1,2,\ldots,K-1)$ می‌باشد، خود نیز یک متغیر تصادفی است. بنابراین می‌توان امید ریاضی آن را به عنوان مقداری برای *P-value* در نظر گرفت.

$$P-value=1-E(H_{(K-1,N-K)}(A))$$

برای محاسبه *P-value* در رابطه بالا می‌توان نمونه‌هایی تصادفی با حجم زیاد (۱۰۰۰) از توزیع بتا با درجات آزادی مربوط به B_j ها انتخاب و مقادیر آن را در رابطه (٤) قرار داد. در این صورت تعداد مقادیر بدست آمده برای $1-H_{(K-1,N-K)}(A)$ برابر حجم نمونه‌ها خواهد بود. اگر میانگین این مقادیر بجای $1-E(H_{(K-1,N-K)}(A))$ قرار داده شود، مقدار *P-value* بدست خواهد آمد. در نهایت فرض H_0 در مقابل فرض H_1 در سطح معنی‌داری α رد خواهد شد؛ اگر $P-value\le\alpha$.

۳-۲ روش چن (۱۹۹۸)

مجدداً مسأله مقایسه میانگینهای K جامعه نرمال با واریانسهای نابرابر را در نظر بگیرید. فرض کنید که

$$X_{ij}\sim N(\mu_i,\sigma_i^2)\qquad i=1,2,\ldots,K\qquad j=1,2,\ldots,n_i$$

چن[5] در سال ۱۹۹۸ روش جدیدی را برای مقایسه میانگینها ارائه داد[۳]. بدین ترتیب که از هر نمونهٔ بدست آمده n_i-1 مشاهده را به دلخواه (یا به تصادف) انتخاب می‌کنیم. میانگین و واریانس نمونه‌های جدید عبارت خواهد بود از

$$\overline{X}_i^*=\sum_{j=1}^{n_i-1}\frac{X_{ij}}{n_i-1},\quad S_i^{*2}=\sum_{j=1}^{n_i-1}\frac{(X_{ij}-\tilde{X}_i^*)^2}{n_i-2}$$

اکنون وزنهای زیر را در نظر می‌گیریم

$$U_i=\frac{1}{n_i}+\frac{1}{n_i}\sqrt{\frac{1}{n_i-1}\left(\frac{S_{\max}^2}{S_i^{*2}}-1\right)}$$

$$V_i=\frac{1}{n_i}-\frac{1}{n_i}\sqrt{(n_i-1)\left(\frac{S_{\max}^2}{S_i^{*2}}-1\right)}$$

که $S_{\max}^2=\max\left(S_1^{*2},S_2^{*2},\ldots,S_K^{*2}\right)$. میانگینهای وزنی بر اساس وزنهای بالا به صورت $\tilde{X}_i=\sum_{j=1}^{n_i}W_{ij}X_{ij}$ تعریف می‌شوند که در آن W_{ij} عبارتست از

$$W_{ij}=\begin{cases}U_i & \text{اگر } 1\le j\le n_i-1\text{،}\\ V_i & \text{اگر } j=n_i\text{،}\end{cases}$$

آماره $\tilde{F}$ را بصورت زیر در نظر بگیرید

$$\tilde{F}=\sum_{i=1}^{K}\left(\frac{\tilde{X}_i-\overline{\tilde{X}}-(\mu_i-\overline{\mu})}{\sqrt{S_{\max}^2/n_i}}\right)^2$$

[4] Weerahandi

[5] Chen

در مواردی که $n_1 = n_2 = \ldots = n_K = n$ باشد، بر پایه یافته‌های بیشاپ و همکارانش[۶] (۱۹۷۸) توزیع تقریبی زیر را برای $\widetilde{F}$ در نظر گرفتند[۱].

$$\widetilde{F} = \frac{n-1}{n-3}\chi^2_{(K-1)}$$

چن با استفاده از روش شبیه‌سازی نشان داد که این آماره آزمون را می‌توان در حالتی که حجم نمونه‌ها مساوی نیستند نیز با درنظر گرفتن $n = \sum_{i=1}^{K} n_i / K$ بکار برد.

با این توضیحات فرض H_0 در مقابل فرض H_1 در سطح معنی‌داری α رد می‌شود اگر تحت فرض H_0 داشته باشیم

$$\widetilde{F} = \sum_{i=1}^{K}\left(\frac{\widetilde{X}_i - \overline{\widetilde{X}} - (\mu_i - \overline{\mu})}{\sqrt{S^2_{\max} / N}}\right)^2 > \frac{n-1}{n-3}\chi^2_{(K-1,1-\alpha)}$$

۲-۴ روش ولچ[۷] (۱۹۵۱)

مبنای روش ولچ، استفاده از روش گشتاورها برای برآورد دومین درجه آزادی آزمون F است که به دلیل بر قرار نبودن شرط تساوی واریانسها باید برآورد شود. در این روش آماره آزمون بصورت زیر بیان شده است.

$$F_W = \frac{\frac{1}{K-1}\sum_{i=1}^{K} g_i(\overline{X}_i - \widetilde{X})^2}{1 + \left(\frac{2(K-2)}{K^2-1}\right)\sum_{i=1}^{K}\frac{\left(1 - g_i / g\right)^2}{n_i - 1}}$$

که در آن

$$S_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2, \quad g_i = n_i / S_i^2$$

$$\widetilde{X} = \sum_{i=1}^{K} g_i \overline{X}_i / g \qquad , \qquad g = \sum_{i=1}^{k} g_i$$

F_W دارای توزیع تقریبی $F_{(K-1,\hat{f}_W)}$ است و $\hat{f}_W$ نیز از رابطه زیر برآورد می‌شود [۷] و[۸].

$$\hat{f}_W = \left(\frac{3}{K^2-1}\sum_{i=1}^{K}\frac{(1-g_i/g)^2}{n_i - 1}\right)^{-1} \tag{۵}$$

با این توضیحات فرض H_0 در مقابل فرض H_1 در سطح معنی‌داری α رد می‌شود اگر $F_W > F_{(K-1,\hat{f}_W,1-\alpha)}$.

۲-۵ روش اسکات ـ اسمیت (۱۹۷۱)

در این قسمت با ایده روش اسکات ـ اسمیت[۸] [۵] که برای مقایسه میانگینهای دو جامعه مطرح شده است، روشی برای مقایسه میانگینهای چند جامعه نرمال با واریانسهای نابرابر پیشنهادمی‌شود. اگر $\mu_1 = \mu_2 = \ldots = \mu_K$ باشد $\mu_i = \overline{\mu}$ (به ازای $i = 1,2,\ldots,K$) خواهد شد. بدیهی است که هر اندازه که این اختلاف بیشتر باشد، انحراف از فرض H_0 بیشتر خواهد شد. بر این اساس می‌توان آماره آزمون فرض H_0 در مقابل H_1 را بصورت زیر بیان کرد.

$$F_S = \sum_{i=1}^{K}\frac{(\overline{X}_i - \overline{X} - (\mu_i - \overline{\mu}))^2}{S_i^{*2} / n_i}$$

که در آن

$$\overline{X} = \frac{\sum_{i=1}^{K} n_i \overline{X}_i}{\sum_{i=1}^{K} n_i} \quad , \quad \overline{\mu} = \frac{\sum_{i=1}^{K} n_i \mu_i}{\sum_{i=1}^{k} n_i}$$

و $S_i^{*2} = \frac{n_i - 1}{n_i - 3} S_i^2$ است. این آماره تحت فرض H_0 بصورت زیر در می‌آید.

$$Fs = \sum_{i=1}^{K}\frac{(\overline{X}_i - \overline{X})^2}{S_i^{*2} / n_i}$$

با توجه به اینکه $\overline{X}_i \sim N(\mu_i, \sigma_i^2 / n_i)$ ، پس

$$T_i = \frac{\overline{X}_i - \mu_i}{\sqrt{S_i^2 / n_i}} \sim t_{(n_i - 1)}$$

بنابراین T_i دارای میانگین صفر و واریانس $\frac{n_i - 1}{n_i - 3}$ می‌باشد و بنابر قضیه حد مرکزی،

---

۶ Bishop et.al.

۷ Welch

۸ Scott – Smith

(به طور جانبی)
$$t_i^* = \frac{\overline{X}_i - \mu_i}{\sqrt{\frac{n_i-1}{n_i-3}S_i^2/n_i}} \sim N(0,1)$$

در نتیجه

$$\sum_{i=1}^{K}\left(\frac{\overline{X}_i - \mu_i}{\sqrt{\frac{n_i-1}{n_i-3}S_i^2/n_i}}\right)^2 \sim \chi^2_{(K)}$$

بنابراین $F_S = \sum_{i=1}^{K} n_i(\overline{X}_i - \overline{X})^2 \big/ S_i^{*2} \sim \chi^2_{(K-1)}$ است. بر این اساس فرض H_0 در مقابل فرض H_1 در سطح معنی‌داری α رد می‌شود اگر $F_S > \chi^2_{(K-1)}$. به دلیل استفاده از قضیه حد مرکزی، این آزمون برای حجم نمونه بیش از ۳۰ قابل استفاده است.

۲-۶ مقایسه روشها

برای مقایسه روشهای پیشنهادی، از طریق روش شبیه‌سازی عمل نموده‌ایم. بدین منظور نمونه‌هایی به حجم $n_i = (5,5,5),(10,10,10),\dots,(5,10,15,20)$ از جوامعی نرمال با میانگینهایی مساوی با $d = \mu_i = (0,0,0),(1,2,3),\dots,(1,3,5,7)$ و واریانسهایی مساوی با $\sigma_i^2 = (1,1,1),(1,2,4),\dots,(30,20,10,1)$ انتخاب شد و این عمل ۱۰۰۰ مرتبه تکرار شد. در هر یک از این ۱۰۰۰ مرتبه، فرض تساوی میانگینهای جوامع از طریق هر یک از روشهای پیشنهاد شده، در سطوح معنی‌داری ۰/۰۵ و $\alpha = 0/01$ آزمون گردید. تعداد دفعاتی که هر یک از روشهای آزمون به ازای مقادیر مختلف σ_i^2, μ_i, n_i ، و α ، فرض H_0 را در مقابل H_1 رد کرده‌اند در جداول (الف-۱) الی (الف-۴) از ضمیمه (الف) آورده شده است. اگر مقادیر متناظر با $\mu_i = (0,0,0),(0,0,0,0)$ را بر ۱۰۰۰ تقسیم کنیم، اندازه واقعی آزمونها (در مقابل اندازه قراردادی[9] آزمونها) بدست خواهد آمد و به ازای سایر مقادیر μ_i ، حاصل این تقسیم توان هر آزمون را بدست خواهد داد. همچنین به منظور ساده‌سازی، روشهای آزمون را با علامتهای اختصاری زیر نشان خواهیم داد.

آزمون F = "F"

آزمون ویراهندی = "We"

آزمون چن = "Ch"

آزمون اسکات-اسمیت = "S"

آزمون ولچ = "W"

۲-۶-۱ سطح معنی‌داری $\alpha = 0/05$

همانگونه که در جداول (الف-۱) و (الف-٤) مشاهده می‌شود، با تغییر عواملی مانند حجم نمونه‌ها و نسبت حجم نمونه‌ها به بزرگی واریانسهای جوامع، رتبه‌بندی آزمونها بر اساس اندازه آزمون عوض می‌شود. بنابراین در این قسمت پس از رتبه‌بندی آزمونها بر اساس توان آزمون به بررسی اندازه آزمونها در شرایط مختلفی که با توجه به عوامل بالا تعیین شده‌اند، می‌پردازیم.

با توجه به جداول (الف-۱) و (الف-٤) می‌توان روشهای مورد بررسی به غیر از آزمون F را به ترتیب Ch ،W ،We و S رتبه‌بندی نمود. بدین ترتیب Ch دارای کمترین توان آزمون و S دارای بیشترین توان آزمون می‌باشد.

حجم نمونه‌ها مساوی باشند: در این شرایط، آزمون در تمام مواردی که واریانسهای جوامع یکسان نیستند، دارای اندازه آزمون بالای ۰/۰۵ است. آزمونهای We و W از نظر اندازه آزمون، نزدیک به هم بوده و نسبت به دو آزمون دیگر از اندازه آزمون کمتری برخوردار هستند. البته در مواردی که حجم نمونه‌ها بالاتر از ۱۰ باشد، Ch نیز از لحاظ اندازه آزمون در کنار آزمونهای We و W قرار می‌گیرد. آزمون S در همه موارد، آزمون Ch تقریباً در همه موارد و آزمونهای We و W در بعضی موارد، دارای اندازه آزمون بیشتر از ۰/۰۵ هستند.

حجم نمونه بزرگتر به جامعه دارای واریانس بیشتر اختصاص داده شود: در این شرایط، آزمون F نسبت به آزمونهای دیگر از اندازه و توان آزمون کمتری برخوردار است. برای مثال در جدول (الف-۱) به ازای $n_i = (5,10,15)$ ، $\mu_i = (0,0,0)$ و $\sigma_i^2 = (1,2,4)$ اندازه آزمون F مساوی ۰/۰۱٦ است که نسبت به آزمونهای دیگر کمترین است. در همین جدول، توان این آزمون به ازای $\mu_i = (1,2,3)$ مساوی ۰/۰۹۸ می‌باشد که باز هم نسبت به آزمونهای دیگر کمترین است. آزمون We دارای کمترین اندازه آزمون بوده و آزمونهای Ch

---

[9] Normal Size

W، S و S به ترتیب در مراتب بعدی قرار دارند. برای مثال در جدول (الف-۲) به ازای $n_i = (۵,۱۰,۱۵,۲۰)$ ، $\mu_i = (۰,۰,۰,۰)$ و $\sigma_i^۲ = (۱,۲,۴,۶)$ اندازه آزمون We مساوی ۰/۰۳۶ و اندازه آزمونهای Ch، W و S به ترتیب برابر با ۰/۰۴۰ ، ۰/۰۴۶ و ۰/۰۷۴ شده است. تمام آزمونها به غیر از S دارای اندازه آزمون کمتر از ۰/۰۵ هستند.

حجم نمونه بزرگتر به جامعه دارای واریانس کمتر اختصاص داده شود: در این شرایط، آزمون F نسبت به سایر آزمونها دارای اندازه بیشتری است. برای مثال در جدول (الف-۱) به ازای $n_i = (۵,۱۰,۱۵)$ ، $\mu_i = (۰,۰,۰)$ و $\sigma_i^۲ = (۴,۲,۱)$ اندازه این آزمون مساوی ۰/۱۹۴ است که نسبت به سایر آزمونها بیشترین است. البته توان آن نسبت به آزمونهای دیگر بیشتراست. آزمون We دارای کمترین اندازه آزمون بوده و آزمونهای W، Ch و S به ترتیب در مراتب بعدی قرار دارند. برای مثال در جدول (الف-۲) به ازای $n_i = (۵,۱۰,۱۵,۲۰)$ ، $\mu_i = (۰,۰,۰,۰)$ و $\sigma_i^۲ = (۶,۴,۲,۱)$ اندازه آزمون We مساوی ۰/۰۵۶ و اندازه آزمونهای W، Ch و S به ترتیب برابر با ۰/۰۵۸، ۰/۰۶۸ و ۰/۰۸۲ شده است. تمام آزمونها دارای اندازه آزمون بالای ۰/۰۵ هستند.

علاوه بر موارد فوق مشاهده می‌شود که اندازه و توان آزمون به ازای تغییر در تفاوت واریانسها، دچار تغییرات فاحشی می‌شود. این موضوع بر استوار نبودن آزمون نسبت به شرط تساوی واریانسها، دلالت می‌کند. برای مثال، در جدول (الف-۱) دامنه تغییرات اندازه آزمون F به ازای $n_i = (۵,۱۰,۱۵)$ از ۰/۰۱۶ تا ۰/۲۷۸ می‌باشد که در مقایسه با سایر آزمونها بیشترین است. در همین جدول به ازای $\mu_i = (۱,۲,۳)$ دامنه تغییرات توان آزمون F از ۰/۰۲۲ تا ۰/۹۶۲ می‌باشد که باز هم در مقایسه با سایر آزمونها بیشترین است.

۲-۷ سطح معنی‌داری $\alpha = ۰/۰۱$

همانگونه که در جداول (الف-۳) و (الف-۴) مشاهده می‌شود، می‌توان همان رتبه‌بندیهای قسمت قبل را برای $\alpha = ۰/۰۱$ نیز در نظر گرفت. از اینرو از ذکر مجدد آنها خودداری می‌شود.

۳. نتیجه‌گیری

با در نظر گرفتن رتبه‌بندیهای ارائه شده در قسمتهای قبل و با توجه به جداول (الف-۱) الی (الف-٤)، در حالت عدم تساوی واریانسها، آزمون در مواردی که دارای اندازه آزمون مناسبی (کمتر از α) است، توانی کمتر از سایر آزمونها دارد.همچنین در مقایسه با سایر آزمونها نسبت به شرط تساوی واریانسها نااستوارتر است. از اینرو این آزمون را مناسب تشخیص نداده، کنار می‌گذایم.

اگر چه آزمون S نسبت به سایر آزمونها دارای توان بالاتری است، ولی چون از نظر اندازه آزمون مقادیری بیشتر از α را به خود اختصاص داده است، آزمون مناسبی به نظر نمی‌رسد. بنابراین این آزمون را نیز کنار می‌گذاریم.

آزمون Ch فقط در مواردی که حجم نمونه‌ها متناسب با بزرگی واریانسهای جوامع هستند، آزمون نسبتاً قابل قبولی است، ولی در این شرایط نیز آزمون We از آزمون مذکور بهتر است. از اینرو آزمون Ch را نیز کنار می‌گذاریم.

با وجود اینکه آزمون W در شرایطی که حجم نمونه‌ها یکسان هستند، آزمون نسبتاً مناسبی به نظر می‌رسد، ولی در همین شرایط اندازه این آزمون نزدیک به آزمون We بوده و توان آن کمتر از آزمون We می‌باشد. بنابراین آزمون W را نیز کنار می‌گذاریم.

بدین ترتیب با حذف آزمونهای S ، Ch و W آزمون We را می‌توان به عنوان آزمون مناسب، توصیه نمود. ولی آزمون مذکور نیز در شرایطی که حجم نمونه‌ها با بزرگی واریانسها نسبت عکس دارند، دارای اندازه آزمون بالاتر از α می‌باشد. علاوه بر این محاسبات مربوط به این آزمون پیچیده و حجیم است. از اینرو در شرایط فوق برای یافتن آزمونی بهتر همچنان باید تحقیق کرد.

نکته: همانگونه که در جداول (الف-۱) تا (الف-٤) مشاهده می‌شود، هنگامی که حجم نمونه بزرگتر به جامعه دارای واریانس بیشتر اختصاص داده می‌شود، انداره آزمونها کوچکتر شده و توان آنها افزایش پیدا می‌کند. از اینرو تعیین حجم نمونه‌ها از اهمیت خاصی برخوردار است.

جدول (الف-۱) اندازه و توان آزمونها به ازای $\alpha = ۰/۰۵$

n_i	μ_i	σ_i^2	F	Wc	Ch	W	S
5,5,5	0,0,0	1,10.20	60	50	92	48	222
		1,2,4	68	40	68	42	82
		1,1,1	50	38	90	40	62
		4,2,1	68	40	68	42	82
		20,10,1	60	50	92	48	222
	1,2,3	1,10,20	90	72	68	62	406
		1,2,4	166	134	128	132	400
		1,1,1	708	624	562	620	646
		4,2,1	166	134	128	132	400
		20,10,1	90	72	78	62	406
	1,3,5	1,10,20	98	94	84	88	744
		1,2,4	418	517	310	502	906
		1,1,1	998	998	974	994	996
		4,2,1	418	517	310	502	906
		20,10,1	98	94	94	88	744
10 ,10 ,10	0,0,0	1,10,20	82	64	60	62	154
		1,2,4	64	52	54	54	92
		1,1,1	50	50	56	52	52
		4,2,1	64	52	54	54	92
		20,10,1	82	64	60	62	154
	1,2,3	1,10,20	74	50	48	48	776
		1,2.4	268	310	164	310	656
		1,1,1	982	958	920	947	976
		4,2,1	268	310	164	310	656
		20,10,1	74	50	48	48	776
	1,3,5	1,10,20	104	102	64	94	834
		1,2,4	784	904	518	892	986
		1,1,1	1000	1000	1000	1000	1000
		4,2,1	784	904	518	892	986
		20,10,1	104	102	64	94	834
5,10,15	0,0,0	1,10,20	24	30	41	46	168
		1,2,4	16	32	38	44	90
		1,1,1	50	48	54	54	46
		4,2,1	194	54	94	58	78
		20,10,1	278	62	92	66	270
	1,2,3	1,10,20	22	68	54	64	568
		1.2.4	98	296	276	288	660
		1,1,1	962	918	916	912	1000
		4,2,1	458	278	260	254	504
		20,10,1	272	68	65	54	548
	1,3,5	1,10,20	34	120	72	116	942
		1,2,4	542	902	730	908	1000
		1,1,1	1000	1000	1000	1000	1000
		4,2,1	878	824	764	772	998
		20,10,1	246	86	126	84	928

جدول (الف-۲) اندازه و توان آزمونها به ازای $\alpha = ۰/۰۵$

n_i	μ_i	σ_i^2	F	We	Ch	W	S
5,5,5,5	0,0,0,0	1,10,20,30	76	66	90	66	206
		1,2,4,6,	56	50	90	48	76
		1,1,1,1,	50	50	78	48	54
		6,4,2,1,	56	50	90	48	76
		30,20,10,1	76	66	90	66	206
	1,2,3,4	1,10,20,30	164	214	132	190	558
		1,2,4,6,	526	516	342	486	670
		1,1,1,1,	974	948	836	934	934
		6,4,2,1,	526	516	342	486	670
		30,20,10,1	164	214	132	190	558
	1,3,5,7	1,10,20,30	476	686	290	592	952
		1,2,4,6,	990	998	834	988	998
		1,1,1,1,	1000	1000	998	1000	1000
		6,4,2,1,	990	998	834	988	998
		30,20,10,1	476	686	290	592	952
10,10,10,10	0,0,0,0	1,10,20,30	68	58	54	56	260
		1,2,4,6,	64	42	44	42	62
		1,1,1,1,	50	50	54	50	60
		6,4,2,1,	64	42	44	42	62
		30,20,10,1	68	58	54	56	260
	1,2,3,4	1,10,20,30	262	408	142	366	840
		1,2,4,6,	906	922	626	918	980
		1,1,1,1,	1000	1000	998	1000	1000
		6,4,2,1,	906	922	626	918	980
		30,20,10,1	262	408	142	366	840
	1,3,5,7	1,10,20,30	1000	1000	944	1000	1000
		1,2,4,6,	1000	1000	990	1000	1000
		1,1,1,1,	1000	1000	1000	1000	1000
		6,4,2,1,	1000	1000	990	1000	1000
		30,20,10,1	1000	1000	944	1000	1000
5,10,15,20	0,0,0,0	1,10,20,30	14	38	45	50	164
		1,2,4,6,	18	36	40	46	74
		1,1,1,1,	50	51	54	56	54
		6,4,2,1,	174	56	68	58	82
		30,20,10,1	224	68	100	74	320
	1,2,3,4	1,10,20,30	118	498	434	510	770
		1,2,4,6,	804	934	936	938	976
		1,1,1,1,	1000	1000	1000	1000	1000
		6,4,2,1,	952	930	901	898	962
		30,20,10,1	474	386	308	296	818
	1,3,5,7	1,10,20,30	696	996	894	992	998
		1,2,4,6,	1000	1000	1000	1000	1000
		1,1,1,1,	1000	1000	1000	1000	1000
		6,4,2,1,	1000	1000	998	1000	1000
		30,20,10,1	934	944	848	908	1000

جدول (الف-۳) اندازه و توان آزمونها به ازای $\alpha = ۰/۰۱$

n_i	μ_i	σ_i^2	F	Wc	Ch	W	S
5,5,5	0,0,0	1,10,20	18	12	32	14	108
		1,2,4	12	4	42	4	28
		1,1,1	10	10	42	8	18
		4,2,1	12	4	42	4	28
		20,10,1	18	12	32	14	108
	1,2,3	1,10,20	92	66	46	54	690
		1,2,4	174	156	138	144	422
		1,1,1	718	626	550	606	628
		4,2,1	174	156	138	144	422
		20,10,1	92	66	46	54	690
	1,3,5	1,10,20	86	82	70	78	666
		1,2,4	416	498	302	488	804
		1,1,1	1000	994	974	988	964
		4,2,1	416	498	302	488	804
		20,10,1	86	82	70	78	666
10,10,10	0,0,0	1,10,20	22	11	14	12	174
		1,2,4	11	8	17	9	34
		1,1,1	10	8	15	8	20
		4,2,1	11	8	17	9	24
		20,10,1	22	11	14	12	74
	1,2,3	1,10,20	94	54	46	50	738
		1,2,4	304	350	172	342	620
		1,1,1	972	952	894	948	960
		4,2,1	304	350	172	342	640
		20,10,1	94	54	46	50	738
	1,3,5	1,10,20	128	96	68	82	804
		1,2,4	802	908	528	890	936
		1,1,1	1000	1000	998	1000	1000
		4,2,1	802	908	528	890	976
		20,10,1	128	96	68	82	878
5,10,15	0,0,0	1,10,20	4	6	6	10	78
		1,2,4	6	6	8	10	40
		1,1,1	10	6	10	10	20
		4,2,1	36	4	34	16	24
		20,10,1	66	16	28	18	90
	1,2,3	1,10,20	18	90	78	82	750
		1,2,4	220	302	280	298	662
		1,1,1	816	696	651	682	970
		4,2,1	502	288	254	266	650
		20,10,1	178	50	44	46	779
	1,3,5	1,10,20	150	568	436	560	882
		1,2,4	946	966	910	943	998
		1,1,1	1000	1000	1000	1000	1000
		4,2,1	994	952	898	924	994
		20,10,1	538	286	198	204	890

جدول (الف-٤) اندازه و توان آزمونها به ازای $\alpha = ۰/۰۱$

n_i	μ_i	σ_i^2	F	We	Ch	W	S
5,5,5,5	0,0,0,0	1,10,20,30	8	16	36	18	108
		1,2,4,6,	24	14	48	14	48
		1,1,1,1,	10	6	56	6	22
		6,4,2,1,	24	14	48	14	48
		30,20,10,1	8	16	36	18	108
	1,2,3,4	1,10,20,30	24	178	150	166	412
		1,2,4,6,	526	816	796	802	890
		1,1,1,1,	1000	1000	990	994	996
		6,4,2,1,	526	816	796	802	890
		30,20,10,1	24	178	150	166	412
	1,3,5,7	1,10,20,30	408	938	768	910	956
		1,2,4,6,	1000	1000	1000	1000	1000
		1,1,1,1,	1000	1000	1000	1000	1000
		6,4,2,1,	1000	1000	1000	1000	1000
		30,20,10,1	408	938	768	910	956
10,10,10,10	0,0,0,0	1,10,20,30	26	14	14	14	98
		1,2,4,6,	12	6	8	8	36
		1,1,1,1,	10	5	10	7	16
		6,4,2,1,	12	6	8	8	36
		30,20,10,1	26	14	14	14	98
	1,2,3,4	1,10,20,30	48	78	58	66	422
		1,2,4,6,	256	242	186	224	514
		1,1,1,1,	854	762	650	702	820
		6,4,2,1,	256	242	186	224	514
		30,20,10,1	48	78	58	66	422
	1,3,5,7	1,10,20,30	580	854	302	796	998
		1,2,4,6,	1000	1000	972	1000	1000
		1,1,1,1,	1000	1000	1000	1000	1000
		6,4,2,1,	1000	1000	972	1000	1000
		30,20,10,1	580	854	302	796	998
5,10,15,20	0,0,0,0	1,10,20,30	4	12	13	14	68
		1,2,4,6,	6	8	10	12	28
		1,1,1,1,	10	12	8	14	24
		6,4,2,1,	82	20	30	26	30
		30,20,10,1	96	24	36	28	160
	1,2,3,4	1,10,20,30	40	220	190	214	632
		1,2,4,6,	512	788	760	778	870
		1,1,1,1,	1000	996	971	989	1000
		6,4,2,1,	870	746	558	656	910
		30,20,10,1	306	182	114	128	744
	1,3,5,7	1,10,20,30	420	934	764	910	992
		1,2,4,6,	1000	1000	1000	1000	1000
		1,1,1,1,	1000	1000	1000	1000	1000
		6,4,2,1,	1000	1000	996	1000	1000
		30,20,10,1	824	756	594	614	998

مراجع

[1] Bishop, T.A., Dudewicz, E.J., Juritz, J.M. and Stephan, M.A., 1978. *Percentage Points of a Quadratic Form in Student t Variates*, Biometrika, Vol.65, pp. 435- 439.

[2] Buning, H., 1995. *Robust ANOVA, the Case of Two-Sided Alternatives,* Proccedings of International Conference on Statistical Methods and Statistical Computing for Quality and Productivity Improvement, August 17- 19, Seoul, Korea.

[3] Chen, S.Y. and Chen, H.J., 1998. *Single – Stage Analysis of Variance Under Hetroscedasticity,* Comm. Stat. Simulation and Computation, Vol.27(3), pp. 641- 666.

[4] Fisher, R.A., 1956. *New Tables of Behrens' test of Significance,* Jour. Roy. Stat. Ser. B, Vol.18, pp. 212- 216.

[5] Scott, A.J. and Smith, T.M.F., 1971. *Interval Estimates for Linear Combinations of Means,* Applied Statistics, Vol.20, pp. 276- 285.

[6] Weerahandi, S., 1995. *Exact Statistical Methods for Data Analysis,* New York, Springer – Verlag.

[7] Welch, B.L., 1947. *The Generalization of Student's Problem When Several Different Population Variances are Involved*, Biometrika, Vol.34, pp. 28- 35.

[8] Welch, B.L., 1951. *On the Comparison of Several Means Values: An Alternative Approach*, Biometrika, Vol.38, pp. 330- 336.

قانون نرمال خطا در تجزیهٔ انسانی، به عنوان یکی از وسیع‌ترین تعمیم‌های طبیعی جلوه می‌کند.

سرفرانسیس گالتون