

آزمون‌های کلاسیک و بوت‌استرپ برابری میانگین‌ها

نصراله ایران پناه*، سمانه نوری امامزاده؛ دانشگاه اصفهان

چکیده

در روش بوت‌استرپ پارامتری با استفاده از باز نمونه‌گیری، آزمون پیشنهادی بوت‌استرپ برای برآورد p - مقدار آماره آزمون مورد نظر ارائه می‌شود. در این مقاله ابتدا آزمون‌های فیشر، کاکران، ولج، جیمز، براون و فرسیت، تقریب فیشر، ویراهاندی و ولج تعدیل یافته و همچنین آزمون بوت‌استرپ پارامتری برای بررسی فرضیه برابری میانگین‌ها ارائه می‌شود. سپس با بررسی‌های شبیه‌سازی برای تعداد تیمارهای مختلف و اندازه نمونه‌های متفاوت، اندازه و توان هر یک از آزمون‌ها محاسبه و با یکدیگر مقایسه می‌شوند. در نهایت برای بررسی عوامل مؤثر بر مقاومت فشاری بتن، اندازه این آزمون‌ها برای داده‌های شرکت سیمان سپاهان محاسبه می‌شود.

مقدمه

طرح آزمایش‌ها در علوم مختلف مانند کشاورزی، زیست‌شناسی، پزشکی و مهندسی استفاده می‌شود. آزمون برابری میانگین‌ها در طرح آزمایش‌ها و همچنین کنترل کیفیت، اهمیت زیادی دارد. در طرح آزمایش‌ها برای بررسی فرضیه برابری میانگین‌های k تیمار با فرض برابری واریانس تیمارها، از آزمون فیشر استفاده می‌شود. حال اگر واریانس‌های تیمارها برابر نباشند، آزمون فیشر برای بررسی فرضیه برابری میانگین‌ها قابل استفاده نیست. این مسئله برای حالت خاص دو تیمار مسئله بهرنز- فیشر نام گرفته است. کاکران [۳] آماره‌ای را بر اساس تعمیم مسئله بهرنز- فیشر برای k جامعه ارائه کرد که این آماره از توزیع تقریبی کای- دو تبعیت می‌کند. ولج [۱۰] تعدیل یافته آزمون کاکران را به عنوان آزمونی برای برابری میانگین‌های k تیمار پیشنهاد کرد که این آماره آزمون از توزیع تقریبی فیشر تبعیت می‌کند. جیمز [۵] مقدار بحرانی را برای آزمون برابری میانگین‌های k تیمار تعیین کرد. براون و فرسیت [۲] آماره‌ای را برای آزمون برابری میانگین‌های k تیمار، پیشنهاد کردند که این آماره تعدیل شده آماره آزمون فیشر است. آسیریو و گورلاند [۱] تقریب آزمون فیشر را برای برابری میانگین‌های k تیمار با واریانس‌های نابرابر ارائه کردند. ویراهاندی [۹] آزمونی را که تعمیمی از آزمون فیشر است، بر اساس الگوریتمی برای محاسبه p - مقدار فرضیه برابری میانگین‌های k تیمار با استفاده از شبیه‌سازی پیشنهاد کرد. مهرتورا [۷] آماره‌ای را برای مسئله بهرنز - فیشر برای k تیمار پیشنهاد کرد که این آماره تعدیل شده آزمون براون و فرسیت است. هارتانگ و همکاران [۴] تعدیل یافته آزمون ولج را ارائه کردند که این آماره

واژه‌های کلیدی: تحلیل واریانس، مسئله بهرنز - فیشر، بوت‌استرپ پارامتری، شبیه‌سازی مونت کارلو.

پذیرش ۹۲/۱۱/۱

دریافت ۹۱/۸/۲۹

*نویسنده مسنول iranpanah@sci.ui.ac.ir

از توزیع تقریبی فیشر تبعیت می‌کند. کریشنامورتی و همکاران [۶] بر اساس آزمون کاکران، روش بوت‌استرپ پارامتری را در تحلیل واریانس پیشنهاد کردند. پارک و پارک [۸] در پژوهشی شبیه‌سازی مقایسه‌ای بین برخی آزمون‌های معرفی شده برای برابری میانگین‌ها با تعداد تیمار زیاد، انجام دادند.

در این مقاله ابتدا آزمون‌های کلاسیک فیشر، کاکران، ولج، جیمز، براون و فرسیت، تقریب فیشر، ویراهاندی و ولج تعدیل یافته برای آزمون برابری میانگین‌ها معرفی می‌شود. سپس آزمون بوت‌استرپ پارامتری برای بررسی فرضیه برابری میانگین‌های k تیمار ارائه می‌شود. در ادامه با بررسی شبیه‌سازی، اندازه و توان آزمون‌های معرفی شده از نظر عددی مقایسه می‌شوند. در انتها برای بررسی عوامل مؤثر بر مقاومت فشاری بتن، اندازه این آزمون‌ها بر روی داده‌های شرکت سیمان سپاهان محاسبه می‌شود. برنامه‌های مورد نیاز برای پژوهش‌های شبیه‌سازی و همچنین تحلیل داده‌های واقعی با استفاده از نرم‌افزار آماری R نوشته شده است.

آزمون‌های کلاسیک برابری میانگین‌ها

فرض کنید جامعه‌ای شامل k تیمار با اندازه نمونه‌های متفاوت باشد. با نمونه‌گیری به اندازه n_i از هر تیمار، مشاهدات نمونه‌ای هر تیمار به صورت $\{X_{ij}; i=1, \dots, k; j=1, \dots, n_i\}$ به دست می‌آید. گیریم X_{i1}, \dots, X_{in_i} نمونه‌ای تصادفی از $N(\mu_i, \sigma_i^2)$ و \bar{X}_i و S_i^2 به ترتیب میانگین و واریانس تیمار i -ام نمونه باشند. هدف آزمون فرضیه $H_0: \mu_1 = \dots = \mu_k$ در مقابل $H_1: \mu_i \neq \mu_j$ برای حداقل یک i و $j (i \neq j; i, j = 1, \dots, k)$ است. اگر σ_i^2 ها معلوم باشند، یک آماره آزمون معمول برای بررسی این فرضیه به صورت

$$T \equiv T(\bar{X}_1, \dots, \bar{X}_k; \sigma_1^2, \dots, \sigma_k^2) = \sum_{i=1}^k \frac{n_i}{\sigma_i^2} \bar{X}_i - \frac{\{\sum_{i=1}^k n_i \bar{X}_i / \sigma_i^2\}^2}{\sum_{i=1}^k n_i / \sigma_i^2} \quad (1)$$

است، که تحت فرضیه صفر $T \sim \chi_{k-1}^2$. بنا بر این فرضیه برابری میانگین‌های k تیمار در سطح a رد می‌شود اگر $P(T > t) < a$ ، که در آن t مقدار مشاهده شده آماره آزمون T است.

در زیر ۸ آزمون کلاسیک برای بررسی فرض برابری میانگین‌های k تیمار ارائه می‌شود:

۱. **آزمون فیشر:** فیشر آماره آزمونی را برای بررسی فرضیه برابری میانگین‌های k تیمار، در حالتی که

واریانس‌های تیمارها نامعلوم و برابر باشند، به صورت

$$F = \frac{(N-k) \sum_{i=1}^k n_i (\bar{X}_i - \bar{X}_{..})^2}{(k-1) \sum_{i=1}^k (n_i - 1) S_i^2} \quad (2)$$

پیشنهاد کرد، که در آن $N = \sum_{i=1}^k n_i$ و $\bar{X}_{..} = \sum_{i=1}^k n_i \bar{X}_i / N$ و تحت فرضیه صفر این آماره از توزیع $F_{k-1, N-k}$ تبعیت می‌کند. بنا بر این فرضیه برابری میانگین‌های k تیمار در سطح a رد می‌شود اگر $P(F > f) < a$ ، که در آن f مقدار مشاهده شده آماره آزمون F است.

۲. **آزمون کاکران:** کاکران [۳] بر اساس آماره T در رابطه (۱)، آماره آزمونی را برای بررسی فرضیه

برابری میانگین‌های k تیمار، در حالتی که واریانس‌های تیمارها نامعلوم و نابرابر باشند، به صورت

$$T_1 \equiv T(\bar{X}_1, \dots, \bar{X}_k; S_1^\tau, \dots, S_k^\tau) = \sum_{i=1}^k \frac{n_i}{S_i^\tau} \bar{X}_i^\tau - \frac{\{\sum_{i=1}^k n_i \bar{X}_i / S_i^\tau\}^\tau}{\sum_{i=1}^k n_i / S_i^\tau} \quad (3)$$

پیشنهاد کرد، که تحت فرضیه صفر از توزیع تقریبی χ_{k-1}^τ تبعیت می‌کند. بنا بر این فرضیه برابری میانگین‌های k تیمار رد می‌شود اگر $P(T_1 > t_1) < a$ ، که در آن مقدار مشاهده شده آماره آزمون T_1 است.

۳. آزمون ولج: ولج [۱۰] آماره آزمونی برای فرضیه برابری میانگین‌های k تیمار بر اساس آماره آزمون

کاکران (۳) به صورت

$$W = \frac{T_1 / k - 1}{1 + (\tau(k - \tau)Q / (k^\tau - 1))}$$

ارائه کرد، که در آن $W_i = n_i / S_i^\tau$ و $Q = \sum_{i=1}^k \frac{1}{n_i - 1} (1 - \frac{W_i}{\sum_{j=1}^k W_j})^\tau$ و تحت فرضیه صفر برای اندازه نمونه‌های

بزرگ از توزیع تقریبی F_{k-1, ν_τ} تبعیت می‌کند، که در آن $\nu_\tau = \frac{(k^\tau - 1)}{3q}$ و q مقدار مشاهده شده Q است.

فرضیه برابری میانگین‌های k تیمار در سطح a رد می‌شود اگر $P(W > w) < a$ ، که W مقدار مشاهده شده آماره آزمون W است. آماره آزمون ولج برای حالت خاص دو تیمار به صورت

$$W = \frac{(\bar{X}_1 - \bar{X}_\tau)^\tau}{\frac{S_1^\tau}{n_1} + \frac{S_\tau^\tau}{n_\tau}} F_{1, \nu_\tau}$$

تبدیل می‌شود، که در آن $r = \frac{S_1^\tau / n_1}{S_\tau^\tau / n_\tau}$ و $\nu_\tau = (1 + r)^\tau \left(\frac{r^\tau}{n_1 - 1} + \frac{1}{n_\tau - 1} \right)^{-1}$

۴. آزمون جیمز: جیمز [۵] یک تقریب مرتبه دوم برای توزیع آماره T_1 ارائه کرد. با فرض

$$c = \frac{1}{\tau} (\tau c_\tau + c_1)_q + \left\{ \frac{1}{16} (\tau c_\tau + c_1)^\tau \left(1 - \frac{k - \tau}{\chi_{k-1, a}^\tau - 1} \right) q^\tau \right. \\ + \left\{ \frac{1}{\tau} (\tau c_\tau + c_1) [(\lambda r_{\tau\tau} - 1 \cdot r_\tau + \varepsilon r_\tau - \tau R_{\tau\tau}^\tau + \lambda R_{\tau\tau} r_\tau - \varepsilon r_\tau^\tau) \right. \\ + (\tau r_{\tau\tau} - \varepsilon r_{\tau\tau} + \tau r_\tau - \tau R_{\tau\tau}^\tau + \varepsilon R_{\tau\tau} r_\tau - \tau r_\tau^\tau) (c_1 + 1) \\ + \frac{1}{\varepsilon} (-r_\tau^\tau + \varepsilon r_\tau r_\tau - \tau r_\tau r_\tau - \tau r_\tau^\tau + \varepsilon r_\tau r_\tau - r_\tau) (\tau c_\tau + \tau c_1 - 1) \\ + (r_{\tau\tau} - \tau r_\tau + \tau r_\tau - r_\tau) (\tau c_\tau + \tau c_1 + c_1) \\ + \frac{\tau}{16} (r_\tau^\tau + \varepsilon r_{\tau\tau} + \tau r_{\tau\tau} - \varepsilon r_\tau + r_\tau) (\tau \tau c_\tau + 1 \tau c_\tau + \tau c_1) \\ + \frac{1}{16} (\tau r_{\tau\tau} + \varepsilon r_\tau - r_\tau) (\tau \tau c_\tau + 1 \tau c_\tau + \tau c_1) \\ \left. \left. + \frac{1}{\varepsilon} (-r_{\tau\tau} + r_\tau^\tau) (\tau \tau c_\tau + \tau c_\tau + c_1) + \frac{1}{\varepsilon} (r_{\tau\tau} - r_\tau r_\tau) (\varepsilon \tau c_\tau + \tau c_\tau + \tau c_1 + \tau c_1) \right\} \right\}$$

که در آن $c_s = \frac{(\chi_{k-1}^s)^s}{(k-1)(k+1)\dots(k+2s-1)}$ ، $r_{st} = \sum_{i=1}^k \frac{1}{(n_i-1)^s} \left(\frac{w_i}{\sum_{i=1}^k w_j} \right)^t$ و q و w_i مقادیر مشاهده

شده Q و W_i در آزمون ولج هستند، مقدار بحرانی این آماره آزمون که تابعی از S_i^r هاست به صورت

$$J_a = \chi_{k-1, a}^r + C + O((n_i-1)^{-r})$$

است. فرضیه برابری میانگین‌های k تیمار در سطح a رد می‌شود، اگر $P(T_1 > t_1 - c) < a$ ، که در آن c مقدار مشاهده شده C در رابطه فوق و t_1 مقدار مشاهده شده آماره آزمون T_1 در رابطه (۳) است.

۵. آزمون براون و فرسیت: براون و فرسیت [۲] آماره آزمون (۴) را که تعدیل شده آماره آزمون فیشر (۲)

است، برای فرضیه برابری میانگین‌های k تیمار به صورت

$$T_{BF} = \frac{\sum_{i=1}^k n_i (\bar{X}_i - \bar{X}_{..})^r}{\sum_{i=1}^k n_i \left(1 - \frac{n_i}{N}\right)^r S_i^r} \quad (۴)$$

پیشنهاد کردند. این آماره تحت فرضیه صفر برای اندازه نمونه‌های بزرگ از توزیع تقریبی $F_{k-1, \nu}$ پیروی می‌کند

که در آن $\nu = \frac{\sum_{i=1}^k \left(1 - \frac{n_i}{N}\right) S_i^r}{\sum_{i=1}^k \left(1 - \frac{n_i}{N}\right)^r S_i^r / (n_i - 1)}$ فرضیه برابری میانگین‌های k تیمار در سطح α رد می‌گردد اگر،

$P(T_{BF} > t_{BF}) < \alpha$ که در آن t_{BF} مقدار مشاهده شده آماره آزمون T_{BF} است.

۶. تقریب آزمون فیشر: آسیریو و گورلاند [۱] تقریب آزمون فیشر را برای بررسی فرضیه برابری

میانگین‌های k تیمار ارائه کردند. آن‌ها نشان دادند تحت فرضیه صفر و برای اندازه نمونه‌های بزرگ،

آماره آزمون فیشر (۲) از توزیع تقریبی F_{ν_1, ν_2} پیروی می‌کند، که در آن

$$\nu_1 = \frac{\left(\sum_{i=1}^k (n_i - 1) S_i^r\right)^r}{\sum_{i=1}^k (n_i - 1) S_i^r} \quad \text{و} \quad \nu_2 = \frac{\left(\sum_{i=1}^k \left(1 - \frac{n_i}{N}\right) S_i^r\right)^r}{\sum_{i=1}^k S_i^r + \left(\sum_{i=1}^k n_i S_i^r / N\right)^r - 2 \sum_{i=1}^k n_i S_i^r / N}$$

برابری میانگین‌های k تیمار در سطح a رد می‌گردد اگر $P(F > \hat{c} f) < \alpha$ ، که در آن f مقدار مشاهده

شده آماره آزمون فیشر در رابطه (۲) و $\hat{c} = \frac{(N-k) \sum_{i=1}^k (N-n_i) S_i^r}{N(k-1) \sum_{i=1}^k (n_i-1) S_i^r}$ هستند.

۷. آزمون ویراهاندی: ویراهاندی [۹] تعمیم آماره آزمون T در رابطه (۱) را برای فرضیه برابری

میانگین‌های k تیمار به صورت

$$T_{GT} = \frac{T}{T_r}; T_r \equiv T_r(\bar{x}_1, \dots, \bar{x}_k; \frac{v_1^r}{u_1}, \dots, \frac{v_k^r}{u_k}) = \sum_{i=1}^k \frac{n_i u_i \bar{x}_i}{v_i^r} - \frac{\left[\sum_{i=1}^k \frac{n_i u_i \bar{x}_i}{v_i^r} \right]^2}{\sum_{i=1}^k \frac{n_i u_i}{v_i^r}} \quad (۵)$$

پیشنهاد کرد، که در آن $v_i^* = (n_i - 1)s_i^2$ ، $U_i \sim \chi_{n_i}^2 - 1$ و T در رابطه (۱) داده شده است. با توجه به این‌که مقدار

آماره T به ازای $X = x$ که در آن مقدار مشاهده شده X است، برابر $\sum_{i=1}^k \frac{n_i \bar{x}_i}{\sigma_i^2} - \frac{\{\sum_{i=1}^k n_i \bar{x}_i / \sigma_i^2\}^2}{\sum_{i=1}^k n_i / \sigma_i^2}$ است

و از طرفی $U_i = \frac{v_i^*}{\sigma_i^2}$ ، در نتیجه مقدار آماره آزمون T_{GT} به‌ازای $X = x$ برابر با یک خواهد شد، بنا بر این

فرضیه صفر در سطح a رد می‌شود اگر $P(T_{GT} > 1) < a$.

الگوریتم پیشنهادی و پراهمندی برای محاسبه p -مقدار با استفاده از شبیه‌سازی مونت کارلو بدین‌صورت است:

۱. تولید متغیر تصادفی T از توزیع χ_{k-1}^2 .
۲. تولید متغیرهای تصادفی U_i به‌طور مستقل از توزیع χ_{k-1}^2 ($i = 1, \dots, k$).

۳. محاسبه آماره T_T از رابطه (۵).

۴. محاسبه آماره آزمون $T_{GT} = \frac{T}{T_T}$.

۵. تکرار m بار مراحل ۱ الی ۴.

۶. محاسبه $P_value = \frac{\#\{T_{GT} > 1\}}{m}$ ، که در آن نماد $\#$ به معنی تعداد است.

۷. رد فرضیه برابری میانگین‌های k تیمار اگر $p_value < \alpha$.

۸. **آزمون ولج تعدیل یافته:** هارتانگ و همکاران [۴] تعدیل شده آزمون ولج را برای بررسی فرضیه برابری میانگین‌های k تیمار به‌صورت

$$T_{MW} = \frac{T / k - 1}{1 + (2(k-2)Q / (k^2 - 1))}$$

پیشنهاد کردند که در آن $T = \sum_{i=1}^k (W_i \bar{X}_i) - \frac{\{\sum_{i=1}^k W_i \bar{X}_i\}^2}{\sum_{i=1}^k W_i}$ ، $Q = \sum_{i=1}^k \frac{1}{n_i - 1} \left(1 - \frac{W_i}{\sum_{i=1}^k W_j} \right)^2$

و $W_i = n_i / (c_i s_i^2)$ و $c_i = \frac{(n_i + 2)}{(n_i + 1)}$. آماره آزمون T_{MW} تحت فرضیه برابری میانگین‌های k تیمار برای

اندازه نمونه‌های بزرگ از توزیع تقریبی $F_{k-1, v}$ تبعیت می‌کند که در آن $v = \frac{(k^2 - 1)}{3q}$ و \bar{q} مقدار مشاهده شده

\bar{Q} است. فرضیه صفر در سطح α رد می‌گردد اگر $P(T_{MW} > t_{MW}) < \alpha$ که t_{MW} مقدار مشاهده شده آماره آزمون T_{MW} است.

آزمون بوت‌استرپ پارامتری

روش بوت‌استرپ پارامتری شامل باز نمونه‌گیری از مدل‌هایی است که پارامترهای آن‌ها از نمونه مشاهده شده برآورده شده‌اند. کریشنامورتی و همکاران [۶] آماره آزمون بوت‌استرپ پارامتری را براساس آماره آزمون T_1 ارائه کردند. به‌علت این‌که آماره آزمون T_1 پایای در مکان است، بدون از دست دادن کلیت مسئله تحت فرضیه برابری میانگین‌های k تیمار فرض کنید $\mu_1 = \dots = \mu_k = 0$. نسخه بوت‌استرپ آماره آزمون T_1 به‌صورت

$$T^* \equiv T^*(\bar{X}_1^*, \dots, \bar{X}_k^*; S_1^{*2}, \dots, S_k^{*2}) = \sum_{i=1}^k \frac{n_i}{S_i^{*2}} \bar{X}_i^* - \frac{\{\sum_{i=1}^k n_i \bar{X}_i^* / S_i^{*2}\}^2}{\sum_{i=1}^k n_i / S_i^{*2}}$$

معرفی می‌گردد، که در آن $\bar{X}_i^* \sim N(0, S_i^{*2} / n_i)$ و $S_i^{*2} \sim S_i^2 \chi_{n_i-1}^2 / (n_i - 1)$. آماره آزمون بوت‌استرپ T^* به صورت

$$T^* \equiv T^*(Z_i^*, C_i^*, s_i^{*2}) = \sum_{i=1}^k \frac{Z_i^* (n_i - 1)}{C_i^*} - \frac{[\sum_{i=1}^k \sqrt{n_i} Z_i^* (n_i - 1) / \{S_i^2 C_i^*\}]^2}{\sum_{i=1}^k (n_i (n_i - 1) / S_i^2 C_i^*)} \quad (6)$$

ساده می‌شود، که در آن $Z_i^* = \sqrt{n_i} \bar{X}_i^* / s_i \sim (1, 0)$ و $C_i^* = (n_i - 1) S_i^{*2} / s_i^2 \sim \chi_{n_i-1}^2$ فرضیه برابری میانگین‌های k تیمار در سطح a رد می‌گردد اگر $T^* > t_1 < a$ که در آن مقدار مشاهده شده آماره آزمون T_1 در رابطه (۳) است.

- الگوریتم پیشنهادی کربشنامورتی و همکاران [۶] برای آزمون فرضیه برابری میانگین‌های k تیمار به روش بوت‌استرپ پارامتری با استفاده از شبیه‌سازی مونت کارلو برای برآورد p -مقدار بدین صورت است:
۱. تولید k متغیر تصادفی مستقل Z_1^*, \dots, Z_k^* از توزیع نرمال استاندارد.
 ۲. تولید متغیرهای تصادفی مستقل C_i^* از توزیع $\chi_{n_i-1}^2$ ($i = 1, \dots, k$).
 ۳. محاسبه آماره آزمون بوت‌استرپ T^* از رابطه (۶).
 ۴. تکرار B بار مراحل ۱ الی ۳ و محاسبه T_1^*, \dots, T_B^* .
 ۵. محاسبه $p_value = \frac{\#\{T_b^* > t_1; b = 1, \dots, B\}}{B}$ که در آن نماد $\#$ به معنی تعداد است.
 ۶. رد فرضیه برابری میانگین‌های k تیمار اگر $p_value < \alpha$.

بررسی شبیه‌سازی

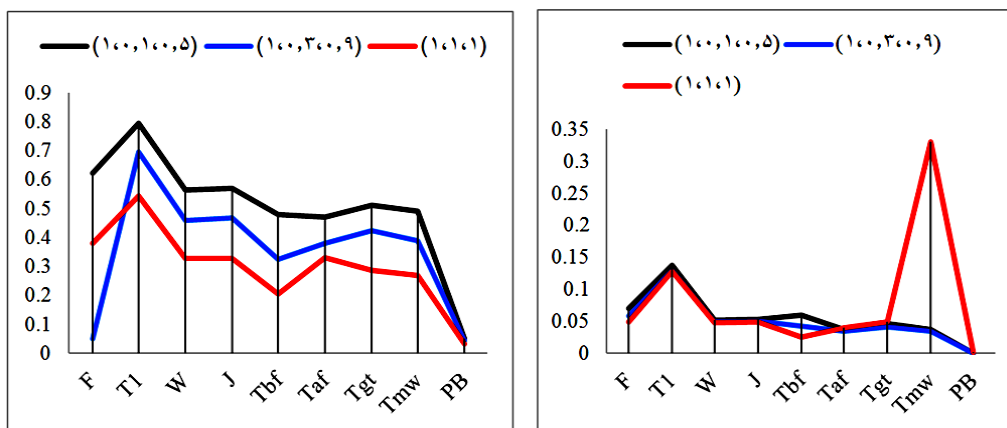
در یک بررسی شبیه‌سازی اندازه و توان آزمون‌های کلاسیک فیشر (F)، کاکران (T_1)، ولج (W)، جیمز (J)، براون و فرسیت (T_{BF})، تقریب فیشر (T_{AF})، ویراهاندی (T_{GT})، ولج تعدیل یافته (T_{MW}) و آزمون بوت‌استرپ پارامتری (PB) برای تعداد تیمارهای ۳، ۶ و ۱۰ و اندازه نمونه‌های متفاوت در سطح $\alpha = 0.05$ به دست آمده است. در این تحقیق تعداد تکرار شبیه‌سازی مونت کارلو ۵۰۰۰ و تعداد تکرار بوت‌استرپ برای آزمون بوت‌استرپ پارامتری ۲۵۰۰ در نظر گرفته شده است. به علت این که آماره آزمون‌های مورد نظر تحت فرضیه صفر، پایایی مکان-مقیاس هستند، بدون از دست دادن کلیت مسئله، اندازه آزمون‌ها تحت فرضیه صفر $\mu_1 = \dots = \mu_k = 0$ و $\sigma_1^2 = 1$ و $0 < \sigma_j^2 \leq 1$ برای $j = 2, \dots, k$ محاسبه می‌شوند. توان آزمون‌ها برای ۳ و ۶ تیمار تحت فرضیه مقابل $\mu_1 = \dots = \mu_{k-2} = 0$ ، $\mu_{k-1} = 1/5$ ، $\mu_k = 1$ و توان آزمون‌ها برای تعداد تیمار ۱۰ تحت فرضیه مقابل $\mu_1 = \dots = \mu_4 = 0$ ، $\mu_5 = 0.2$ ، $\mu_6 = 0.4$ ، $\mu_7 = 0.6$ ، $\mu_8 = 0.8$ ، $\mu_9 = 1$ محاسبه شده است. جدول‌ها و نمودارهای ۱ الی ۳ اندازه و توان این آزمون‌ها را برای ۳ تیمار به ترتیب برای اندازه نمونه‌های (۵، ۵، ۵)، (۱۰، ۱۰، ۱۰) و (۱۵، ۱۰، ۵) و واریانس‌های مختلف نشان می‌دهند. اندازه و توان این آزمون‌ها برای ۶ تیمار برای واریانس‌های مختلف و اندازه نمونه‌های

آزمون‌های کلاسیک و بوت‌استرپ برای میانگین‌ها (۵،۵،۵،۵،۵،۵) و (۱۰،۱۰،۱۰،۱۰،۱۰،۱۰،۱۰) و (۴،۸،۱۲،۲۴،۳۰،۴۰) به ترتیب در جدول‌ها و نمودارهای ۴ الی ۶ نشان داده شده است. جدول‌ها و نمودارهای ۷ الی ۹ اندازه و توان این آزمون‌ها را برای ۱۰ تیمار به ترتیب برای اندازه نمونه‌های (۵،۵،۵،۵،۵،۵،۵،۵،۵،۵)، (۱۵،۱۵،۱۵،۱۵،۱۲،۱۲،۱۲،۴،۴،۴) و (۱۵،۱۷،۱۹،۲۱،۲۳،۲۵،۲۷،۲۹،۳۱،۳۳) و برای واریانس‌های مختلف نشان می‌دهند. لازم به توضیح است که نمودارهای سمت راست اندازه آزمون و نمودارهای سمت چپ توان آزمون را نشان می‌دهند. دلیل استفاده از چنین پارامترها و اندازه نمونه‌ها این است که با انتخاب این پارامترها و اندازه نمونه‌ها، تفاوت اندازه و توان آزمون‌های مختلف بهتر نشان داده می‌شود و مقایسه بین آزمون‌ها آسان‌تر انجام می‌گیرد و اگر نه در انتخاب اندازه نمونه‌ها و واریانس‌ها محدودیت خاصی وجود ندارد.

جدول‌های ۱ الی ۹ نشان می‌دهند با افزایش تعداد تیمارها تنها آزمون‌های ولج، جیمز، ولج تعدیل یافته و بوت‌استرپ پارامتری می‌توانند اندازه آزمون را در سطح $\alpha = 0.05$ کنترل کنند و بقیه آزمون‌ها با افزایش تعداد تیمارها از سطح 0.05 تجاوز می‌کنند. به‌طور مثال اندازه آزمون فیشر برای تعداد تیمار ۱۰ به 0.06 و اندازه آزمون کاکران برای ۶ تیمار به 0.02 می‌رسد. با مقایسه جدول‌های ۱ الی ۶ با جدول‌های ۷ الی ۹ مشاهده می‌شود که هرچه تفاوت میانگین‌ها از فرضیه صفر بیشتر می‌شود، توان آزمون هم بالاتر می‌رود. در جدول ۱ برای سه تیمار و اندازه نمونه (۵،۵،۵) مشاهده می‌گردد وقتی واریانس‌ها برابر باشند، همه آزمون‌ها به جز آزمون کاکران در سطح معمول $\alpha = 0.05$ قرار می‌گیرند و در این میان توان آزمون فیشر از همه بالاتر است. در جدول‌های ۲ و ۳ مشاهده می‌گردد برای ۳ تیمار در حالت نابرابری واریانس‌ها، آزمون‌های ولج و جیمز بیش‌ترین توان را دارند. این در حالی است که در حالت برابری واریانس‌ها آزمون فیشر بیش‌ترین توان را داراست. بر اساس جدول‌های ۴ الی ۹ مشاهده می‌گردد در حالت نابرابری واریانس‌ها برای ۶ و ۱۰ تیمار آزمون بوت‌استرپ پارامتری در بین دیگر آزمون‌هایی که در سطح $\alpha = 0.05$ قرار می‌گیرند، یکی از بهترین آزمون‌ها از نظر ماکسیمم توان است. این در حالی است که در حالت برابری واریانس‌ها و اندازه نمونه‌های نابرابر برای تعداد تیمار ۱۰، آزمون فیشر سطح آزمون را کنترل نمی‌کند و آزمون‌های جیمز و بوت‌استرپ پارامتری و تقریب فیشر پیشنهاد بهترین آزمون هستند.

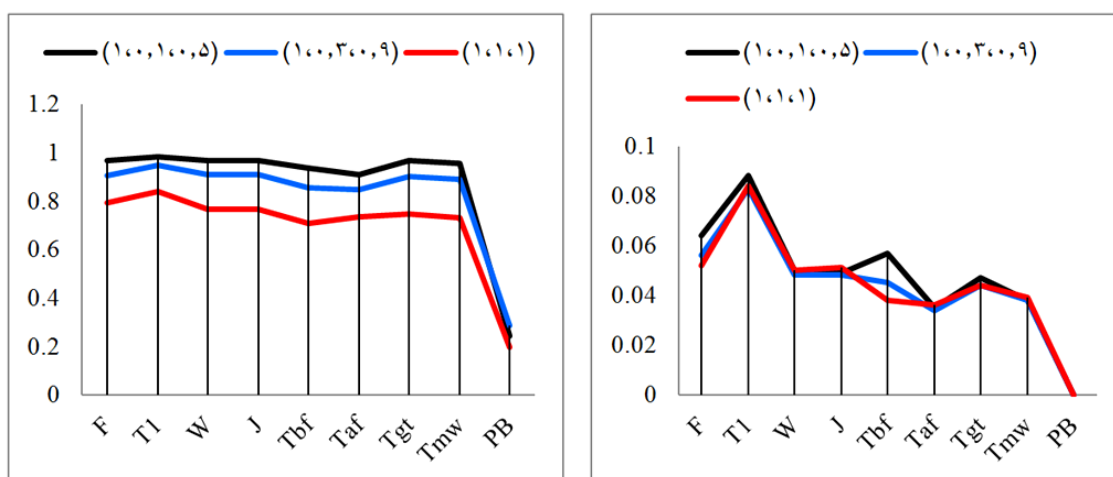
جدول ۱. اندازه و توان آزمون برای ۳ تیمار و اندازه نمونه (۵،۵،۵)

$\sigma_1^2, \sigma_2^2, \sigma_3^2$	(۱،۱،۱)		(۱،۰۰/۳،۰۰/۹)		(۱،۰۰/۱،۰۰/۵)	
	اندازه	توان	اندازه	توان	اندازه	توان
F	۰/۰۴۹	۰/۳۷۹	۰/۰۵۷	۰/۰۴۹۴	۰/۰۶۹	۰/۶۲۲
T_r	۰/۱۲۷	۰/۵۴۱	۰/۱۲۹	۰/۶۹۳	۰/۱۳۷	۰/۷۹۴
W	۰/۰۴۷	۰/۳۲۶	۰/۰۴۸	۰/۴۵۸	۰/۰۵۱	۰/۵۶۳
J	۰/۰۴۹	۰/۳۳۶	۰/۰۵۰	۰/۴۶۶	۰/۰۵۲	۰/۵۶۷
T_{BF}	۰/۰۲۵	۰/۲۰۲	۰/۰۴۲	۰/۳۲۴	۰/۰۵۹	۰/۴۷۸
T_{AF}	۰/۰۳۹	۰/۳۳۰	۰/۰۳۴	۰/۳۷۹	۰/۰۳۷	۰/۴۶۸
T_{GT}	۰/۰۴	۰/۲۸۴	۰/۰۴۰	۰/۴۲۲	۰/۰۴۶	۰/۵۵۱
T_{MW}	۰/۳۳	۰/۲۶۸	۰/۰۳۴	۰/۳۸۶	۰/۰۳۶	۰/۴۹۱
PB	۰/۰۰۱	۰/۰۳۱	۰	۰/۰۴۴	۰	۰/۰۴۹



شکل ۱. اندازه و توان آزمون برای ۳ تیمار و اندازه نمونه (۵،۵،۵)
جدول ۲. اندازه و توان آزمون برای ۳ تیمار و اندازه نمونه (۱۰،۱۰،۱۰)

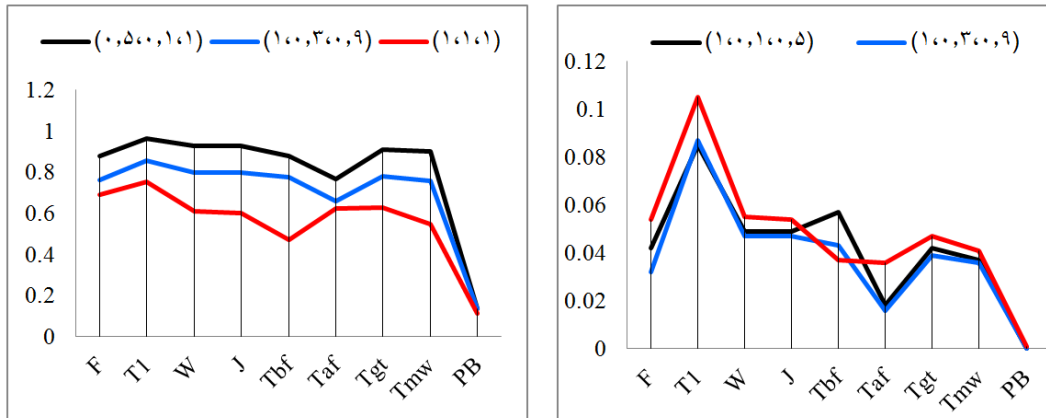
$\sigma_1^2, \sigma_2^2, \sigma_3^2$	(۱۰،۱۰،۱)		(۱۰۰،۳۰۰،۹)		(۱۰۰،۱۰۰،۵)	
آزمون	اندازه	توان	اندازه	توان	اندازه	توان
F	۰/۰۵۲	۰/۷۹۲	۰/۰۵۶	۰/۹۰۰	۰/۰۶۴	۰/۹۶۴
T ₁	۰/۰۸۴	۰/۸۳۸	۰/۰۸۳	۰/۹۴۶	۰/۰۸۸	۰/۹۸۳
W	۰/۰۵۰	۰/۷۶۳	۰/۰۴۸	۰/۹۰۹	۰/۰۰۵	۰/۹۶۵
J	۰/۰۵۱	۰/۷۶۶	۰/۰۴۸	۰/۹۰۹	۰/۰۴۹	۰/۹۶۵
T _{BF}	۰/۰۳۸	۰/۷۰۶	۰/۰۴۵	۰/۸۵۵	۰/۰۵۷	۰/۹۳۶
T _{AF}	۰/۰۳۶	۰/۷۳۲	۰/۰۳۴	۰/۸۴۶	۰/۰۳۵	۰/۹۰۷
T _{GT}	۰/۰۴۴	۰/۷۴۶	۰/۰۴۴	۰/۹۰۱	۰/۰۴۷	۰/۹۶۵
T _{MW}	۰/۰۳۹	۰/۷۳۰	۰/۰۳۸	۰/۸۹۰	۰/۰۳۸	۰/۹۵۳
PB	۰	۰/۱۹۶	۰	۰/۲۸۶	۰	۰/۲۴۲



شکل ۲. اندازه و توان آزمون برای ۳ تیمار و اندازه نمونه (۱۰،۱۰،۱۰)

جدول ۳. اندازه و توان آزمون برای ۳ تیمار و اندازه نمونه (۵،۱۰،۱۵)

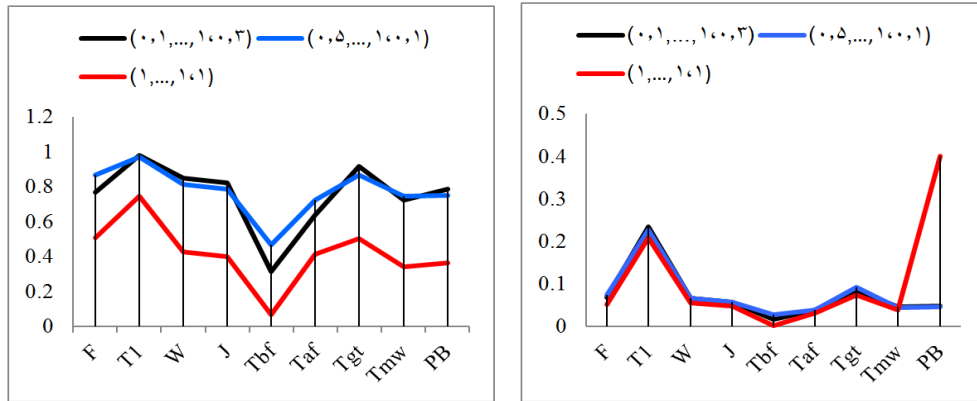
$\sigma_1^2, \sigma_2^2, \sigma_3^2$	(۱،۱،۱)		(۱،۰/۳،۰/۹)		(۱،۰/۱۰،۰/۵)	
آزمون	اندازه	توان	اندازه	توان	اندازه	توان
F	۰/۰۵۴	۰/۶۹۲	۰/۰۳۲	۰/۷۶۲	۰/۰۴۲	۰/۸۷۶
T _l	۰/۱۰۵	۰/۷۵۲	۰/۰۸۷	۰/۸۸۵	۰/۰۸۵	۰/۹۶۲
W	۰/۰۵۵	۰/۶۰۹	۰/۰۴۷	۰/۷۹۷	۰/۰۴۹	۰/۹۲۸
J	۰/۰۵۴	۰/۵۹۹	۰/۰۴۷	۰/۷۹۷	۰/۰۴۹	۰/۹۲۷
T _{BF}	۰/۰۳۷	۰/۴۷۳	۰/۰۴۳	۰/۷۷۶	۰/۰۵۷	۰/۸۷۸
T _{AF}	۰/۰۳۶	۰/۶۲۲	۰/۰۱۶	۰/۶۵۹	۰/۰۱۸	۰/۷۶۵
T _{GT}	۰/۰۴۷	۰/۶۲۷	۰/۰۳۹	۰/۷۷۹	۰/۰۴۲	۰/۹۱۱
T _{MW}	۰/۰۴۱	۰/۵۴۹	۰/۰۳۶	۰/۷۵۶	۰/۰۳۷	۰/۹۰۱
PB	۰/۰۰۱	۰/۱۱۱	۰	۰/۱۳۳	۰	۰/۱۳۶



شکل ۳. اندازه و توان آزمون برای ۳ تیمار و اندازه نمونه (۵،۱۰،۱۵)

جدول ۴. اندازه و توان آزمون برای ۶ تیمار و اندازه نمونه (۵،۵،۵،۵،۵،۵)

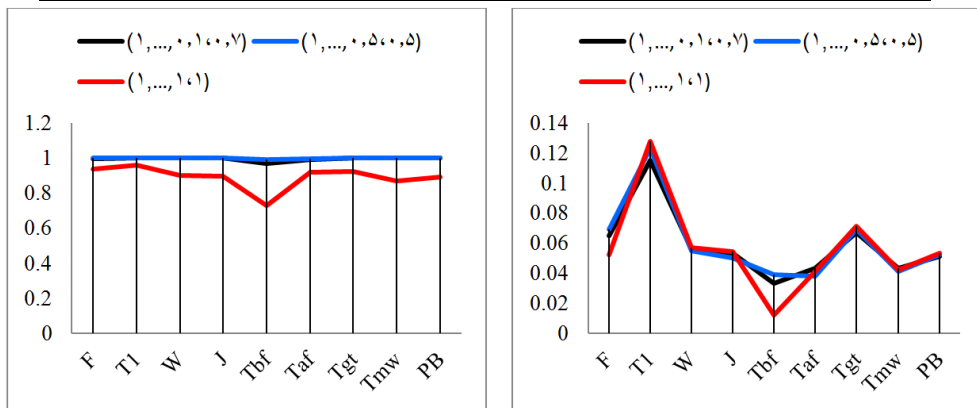
$\sigma_1^2, \dots, \sigma_6^2$	(۱،۱،۱،۱،۱،۱)		(۱،۰/۱،۰/۱،۰/۵،۰/۵،۰/۵)		(۱،۰/۳،۰/۹،۰/۴،۰/۷،۰/۱)	
آزمون	اندازه	توان	اندازه	توان	اندازه	توان
F	۰/۰۵۱	۰/۵۰۶	۰/۰۷۳	۰/۸۶۶	۰/۰۶۷	۰/۷۶۹
T _l	۰/۲۰۸	۰/۷۴۵	۰/۲۲۷	۰/۹۶۷	۰/۲۳۴	۰/۹۸۰
W	۰/۰۵۵	۰/۴۲۵	۰/۰۶۶	۰/۸۱۳	۰/۰۶۵	۰/۸۴۸
J	۰/۰۴۸	۰/۳۹۹	۰/۰۵۷	۰/۷۸۵	۰/۰۵۶	۰/۸۲۰
T _{BF}	۰/۰۰۲	۰/۰۶۶	۰/۰۲۷	۰/۴۶۴	۰/۰۱۶	۰/۳۱۲
T _{AF}	۰/۰۳۰	۰/۴۱۴	۰/۰۳۸	۰/۷۲۲	۰/۰۳۸	۰/۶۳۵
T _{GT}	۰/۰۷۳	۰/۵۰۰	۰/۰۹۱	۰/۸۶۸	۰/۰۸۴	۰/۹۱۳
T _{MW}	۰/۰۳۸	۰/۳۳۸	۰/۰۴۴	۰/۷۴۳	۰/۰۴۶	۰/۷۷۲
PB	۰/۰۴۰	۰/۳۶۳	۰/۰۴۶	۰/۷۵۱	۰/۰۴۸	۰/۷۸۷



شکل ۴. اندازه و توان آزمون برای ۶ تیمار و اندازه نمونه (۰.۵، ۰.۵، ۰.۵، ۰.۵)

جدول ۵. اندازه و توان آزمون برای ۶ تیمار و اندازه نمونه (۱.۰، ۱.۰، ۱.۰، ۱.۰، ۱.۰، ۱.۰)

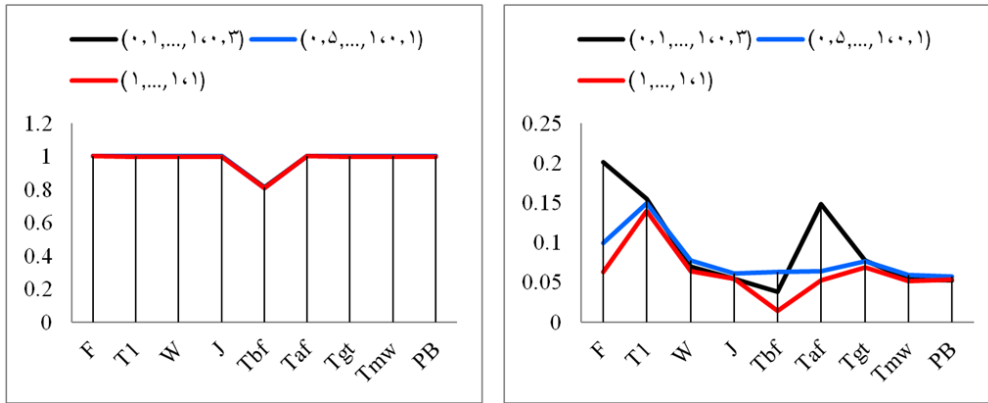
$\sigma_1^2, \dots, \sigma_6^2$	(۱، ۱، ۱، ۱، ۱، ۱)	(۱، ۰/۱، ۰/۱، ۰/۵، ۰/۵، ۰/۵)	(۱، ۰/۳، ۰/۶، ۰/۴، ۰/۷، ۰/۱)
آزمون	اندازه	توان	توان
F	۰/۰۵۲	۰/۹۳۸	۰/۰۶۹
T ₁	۰/۱۲۸	۰/۹۶۰	۰/۱۲۳
W	۰/۰۵۷	۰/۹۰۱	۰/۰۵۵
J	۰/۰۵۴	۰/۸۹۶	۰/۰۵۰
T _{BF}	۰/۰۱۲	۰/۷۲۷	۰/۰۳۹
T _{AF}	۰/۰۴۱	۰/۹۱۹	۰/۰۳۸
T _{GT}	۰/۰۷۱	۰/۹۲۱	۰/۰۶۹
T _{MW}	۰/۰۴۲	۰/۸۶۹	۰/۰۴۱
PB	۰/۰۵۳	۰/۸۹۱	۰/۰۵۲



شکل ۵. اندازه و توان آزمون برای ۶ تیمار و اندازه نمونه (۱.۰، ۱.۰، ۱.۰، ۱.۰، ۱.۰، ۱.۰)

جدول ۶. اندازه و توان آزمون برای ۶ تیمار و اندازه نمونه (۴، ۸، ۱۲، ۲۴، ۳۰، ۴۰)

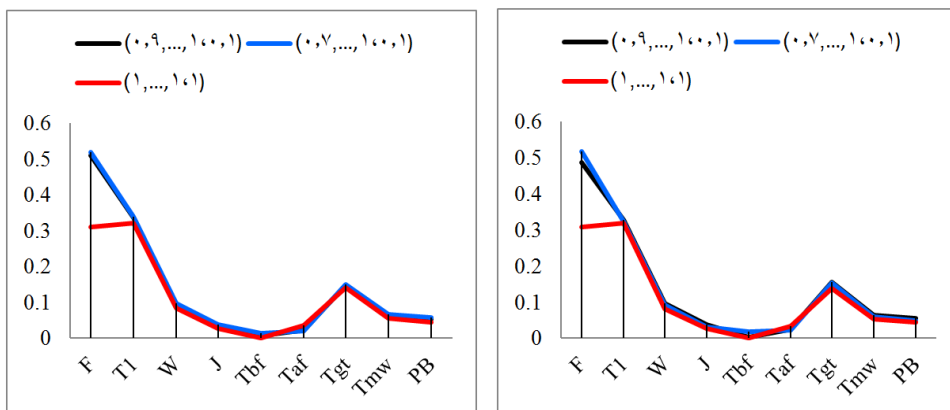
$\sigma_1^2, \dots, \sigma_6^2$	(۱، ۱، ۱، ۱، ۱، ۱)	(۱، ۰/۱، ۰/۱، ۰/۵، ۰/۵، ۰/۵)	(۱، ۰/۳، ۰/۶، ۰/۴، ۰/۷، ۰/۱)
آزمون	اندازه	توان	توان
F	۰/۰۶۳	۱	۰/۰۹۹
T ₁	۰/۱۴۰	۰/۹۹۹	۰/۱۴۹
W	۰/۰۶۴	۰/۹۹۸	۰/۰۷۷
J	۰/۰۵۴	۰/۹۹۸	۰/۰۶۱
T _{BF}	۰/۰۱۴	۰/۸۰۹	۰/۰۶۳
T _{AF}	۰/۰۵۲	۱	۰/۰۶۴
T _{GT}	۰/۰۶۹	۰/۹۹۹	۰/۰۷۶
T _{MW}	۰/۰۵۱	۰/۹۹۸	۰/۰۵۹
PB	۰/۰۵۳	۰/۹۹۷	۰/۰۵۷



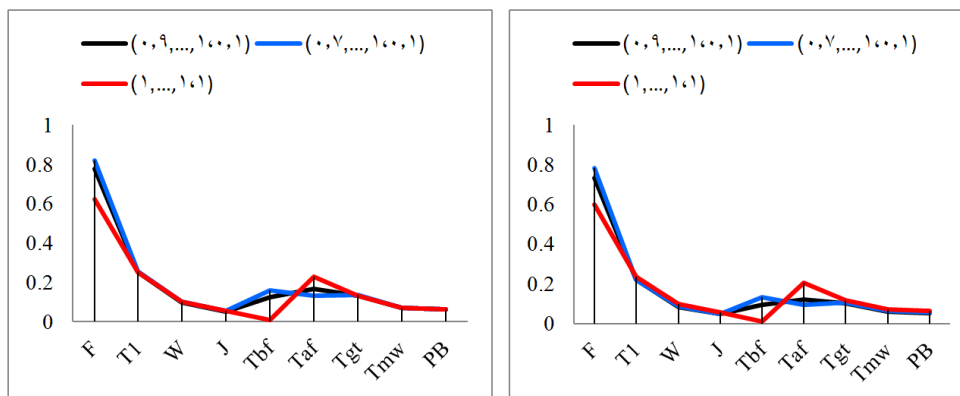
شکل ۶. اندازه و توان آزمون برای ۶ تیمار و اندازه نمونه (۴، ۸، ۱۲، ۲۴، ۳۰، ۴۰)

جدول ۷. اندازه و توان آزمون برای ۱۰ تیمار و اندازه نمونه (۵، ۵، ۵، ۵، ۵، ۵، ۵، ۵، ۵، ۵)

$\sigma_1^2, \dots, \sigma_{10}^2$	(۱، ...، ۱)	(۰.۳، ۰.۳، ۰.۳، ۰.۷، ۰.۷، ۰.۷)	(۰.۴، ۰.۵، ۰.۶، ۰.۷، ۰.۸، ۰.۹)			
آزمون	اندازه	توان	اندازه	توان	اندازه	توان
F	۰/۳۰۷	۰/۳۱۰	۰/۵۱۶	۰/۵۱۹	۰/۴۸۷	۰/۵۱۰
T₁	۰/۳۱۸	۰/۳۲	۰/۳۲۲	۰/۳۳۸	۰/۳۲۷	۰/۳۳۵
W	۰/۰۸۱	۰/۰۸۴	۰/۰۹۱	۰/۰۹۷	۰/۰۹۷	۰/۰۹۰
J	۰/۰۲۶	۰/۰۲۷	۰/۰۳۰	۰/۰۳۷	۰/۰۳۸	۰/۰۳۴
T_{BF}	.	.	۰/۰۱۷	۰/۰۱۴	۰/۰۰۳	۰/۰۰۷
T_{AF}	۰/۰۳۳	۰/۰۳۵	۰/۰۲۲	۰/۰۲۰	۰/۰۲۴	۰/۰۲۳
T_{GT}	۰/۱۳۷	۰/۱۴۱	۰/۱۵۳	۰/۱۵۰	۰/۱۵۴	۰/۱۴۷
T_{MW}	۰/۰۵۳	۰/۰۵۴	۰/۰۵۹	۰/۰۶۶	۰/۰۶۴	۰/۰۶۱
PB	۰/۰۴۴	۰/۰۴۵	۰/۰۴۷	۰/۰۵۶	۰/۰۵۴	۰/۰۵۳



شکل ۷. اندازه و توان آزمون برای ۱۰ تیمار و اندازه نمونه (۵، ۵، ۵، ۵، ۵، ۵، ۵، ۵، ۵، ۵)



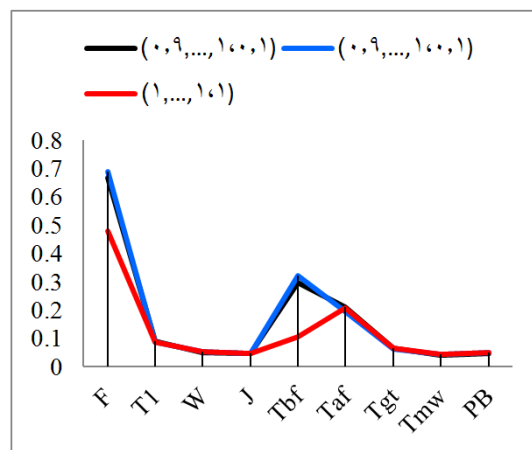
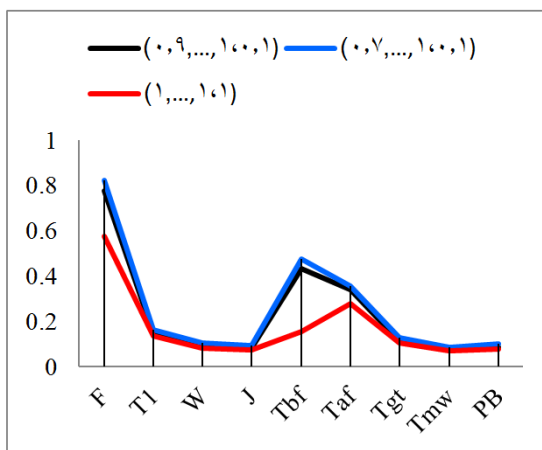
شکل ۸. اندازه و توان آزمون برای ۱۰ تیمار و اندازه نمونه (۴، ۴، ۴، ۱۲، ۱۲، ۱۲، ۱۵، ۱۵، ۱۵، ۱۵)

جدول ۸. اندازه و توان آزمون برای ۱۰ تیمار و اندازه نمونه (۴، ۴، ۴، ۱۲، ۱۲، ۱۲، ۱۵، ۱۵، ۱۵، ۱۵)

$\sigma_1^2, \dots, \sigma_{10}^2$	(۱، ...، ۱)	(۰/۳، ۰/۳، ۰/۳، ۰/۷، ۰/۷، ۰/۷)	(۰/۴، ۰/۵، ۰/۶، ۰/۷، ۰/۸، ۰/۹)	(۰/۳، ۰/۳، ۰/۳، ۰/۷، ۰/۷، ۰/۷)	(۰/۴، ۰/۵، ۰/۶، ۰/۷، ۰/۸، ۰/۹)	
آزمون	اندازه	توان	اندازه	توان	اندازه	توان
F	۰/۵۹۹	۰/۶۲۲	۰/۷۸۵	۰/۸۱۹	۰/۷۳۳	۰/۷۷۹
T _l	۰/۲۳۷	۰/۲۵۳	۰/۲۱۹	۰/۲۵۶	۰/۲۲۰	۰/۲۵۱
W	۰/۰۹۷	۰/۱۰۲	۰/۰۸۶	۰/۱۰۳	۰/۰۸۴	۰/۰۹۷
J	۰/۰۵۶	۰/۰۵۶	۰/۰۴۹	۰/۰۵۴	۰/۰۴۶	۰/۰۵۳
T _{BF}	۰/۰۰۹	۰/۰۰۷	۰/۱۳۲	۰/۱۵۹	۰/۰۹۳	۰/۱۲۳
T _{AF}	۰/۲۰۶	۰/۲۳۰	۰/۰۹۳	۰/۱۳۳	۰/۱۱۹	۰/۱۶۶
T _{GT}	۰/۱۱۷	۰/۱۳۳	۰/۱۰۴	۰/۱۳۷	۰/۱۰۳	۰/۱۳۴
T _{MW}	۰/۰۷	۰/۰۷۱	۰/۰۶۳	۰/۰۷۱	۰/۰۶۰	۰/۰۶۹
PB	۰/۰۶۳	۰/۰۶۱	۰/۰۵۵	۰/۰۶۱	۰/۰۵۳	۰/۰۶۱

جدول ۹. اندازه و توان آزمون برای ۱۰ تیمار و اندازه نمونه (۱۵، ۱۷، ۱۹، ۲۱، ۲۳، ۲۵، ۲۷، ۲۹، ۳۱، ۳۳)

$\sigma_1^2, \dots, \sigma_{10}^2$	(۱، ...، ۱)	(۰/۳، ۰/۳، ۰/۳، ۰/۷، ۰/۷، ۰/۷)	(۰/۴، ۰/۵، ۰/۶، ۰/۷، ۰/۸، ۰/۹)	(۰/۳، ۰/۳، ۰/۳، ۰/۷، ۰/۷، ۰/۷)	(۰/۴، ۰/۵، ۰/۶، ۰/۷، ۰/۸، ۰/۹)	
آزمون	اندازه	توان	اندازه	توان	اندازه	توان
F	۰/۴۷۸	۰/۵۷۶	۰/۶۸۸	۰/۸۲۱	۰/۶۶۷	۰/۷۷۵
T _l	۰/۰۸۷	۰/۱۳۴	۰/۰۸۶	۰/۱۶۱	۰/۰۹۰	۰/۱۴۶
W	۰/۰۵۲	۰/۰۸۱	۰/۰۵۲	۰/۱۰۳	۰/۰۵۱	۰/۰۹۰
J	۰/۰۴۶	۰/۰۷۳	۰/۰۴۶	۰/۰۹۴	۰/۰۴۵	۰/۰۸۲
T _{BF}	۰/۱۰۵	۰/۱۵۴	۰/۳۲۱	۰/۴۷۴	۰/۲۹۶	۰/۴۳۴
T _{AF}	۰/۲۰۸	۰/۲۷۹	۰/۱۹۶	۰/۳۵۴	۰/۲۰۹	۰/۳۴۱
T _{GT}	۰/۰۶۴	۰/۱۰۳	۰/۰۶۲	۰/۱۲۶	۰/۰۶۴	۰/۱۱۳
T _{MW}	۰/۰۴۲	۰/۰۶۸	۰/۰۴۲	۰/۰۸۷	۰/۰۴۰	۰/۰۷۶
PB	۰/۰۵۰	۰/۰۷۸	۰/۰۵۰	۰/۱۰۰	۰/۰۴۷	۰/۰۸۶



شکل ۹. اندازه و توان آزمون برای ۱۰ تیمار و اندازه نمونه (۱۵، ۱۷، ۱۹، ۲۱، ۲۳، ۲۵، ۲۷، ۲۹، ۳۱، ۳۳)

مثال کاربردی

پس از پیدایش بتن همواره سعی بر این بوده است که با شناخت راه‌های مناسب و با ایجاد نسبت‌های معین بین مصالح به‌کار رفته در بتن، و یا با اضافه کردن مواد دیگری به ترکیبات قبلی به‌صورت مواد مضاعف یا مواد جای‌گزینی، مقاومت بتن و نیز سایر خواص مکانیکی و پایایی آن را بهبود بخشند. شرکت سیمان سپاهان اصفهان در پژوهشی برای افزایش مقاومت فشاری بتن‌ها، تأثیر چهار عامل، نوع مصالح سنگی، نسبت آب به مواد سیمانی، درصد میکروسیلیس و حجم الیاف فولادی به‌کار رفته در ساخت بتن را بر روی میزان مقاومت فشاری بتن آزمایش کرده است. در این تحقیق چهار عامل، نوع درشت‌دانه در دو سطح آهکی و کوارتزیتی، عامل نسبت آب به مواد سیمانی در دو سطح ۰/۲۵ و ۰/۴، عامل درصد میکروسیلیس در سه سطح ۰٪ و ۱۰٪ و ۱۵٪ و حجم الیاف فولادی در چهار سطح ۰، ۰/۴، ۰/۸ و ۱/۲ به‌کار برده شده است. اندازه نمونه‌ها در هر سطح عامل با هم برابر هستند. به‌عبارت دیگر اندازه نمونه‌ها در هر سطح عوامل نوع درشت‌دانه و نسبت آب به مواد سیمانی برابر ۷۲، برای هر سطح عامل درصد میکروسیلیس برابر ۴۸ و برای هر سطح عامل حجم الیاف فولادی برابر ۳۶ است. قبل از محاسبه اندازه آزمون‌های مختلف، p- مقدار آزمون‌های کلموگروف- اسمیرنوف لی فورس و شپیرو- ویلک برای بررسی فرضیه نرمال بودن میزان مقاومت فشاری در سطوح مختلف عوامل در جدول ۱۰ ارائه شده است. چنان‌که مشاهده می‌شود داده‌ها در همه سطوح این چهار عامل، در سطح $\alpha = 0/05$ از توزیع نرمال تبعیت می‌کنند. نتایج انجام آزمون‌های فیشر، کاکران، ولج، جیمز، براون و فرسیت، تقریب فیشر، ویراهندی، ولج تعدیل یافته و بوت‌استرپ پارامتری نشان می‌دهند همه آزمون‌ها فرضیه تأثیر نوع درشت‌دانه، نسبت آب به مواد سیمانی و درصد میکروسیلیس را بر روی مقاومت فشاری بتن رد نمی‌شوند. از طرف دیگر همه آزمون‌ها فرضیه برابری میانگین مقاومت فشاری بتن در سطوح مختلف حجم الیاف فولادی را می‌پذیرند و بنا بر این حجم الیاف فولادی تأثیری بر مقاومت فشاری بتن ندارد.

جدول ۱۰. p- مقدار آزمون‌های نرمال بودن میزان مقاومت فشاری سطوح مختلف عوامل

عامل	سطوح	p- مقدار آزمون کلموگروف- اسمیرنوف	p- مقدار آزمون لی فورس	p- مقدار آزمون شپیرو- ویلک
نوع درشت‌دانه	آهکی	۰/۱۶	۰/۸۶	۰/۲۹
	کوارتزیتی	۰/۲۰	۰/۱۸۵	۰/۰۳
نسبت آب به مواد سیمانی	۰/۲۵	۰/۲۰	۰/۶۰	۰/۱۲
	۰/۴	۰/۱۹	۰/۱۸۷	۰/۳۸
درصد میکروسیلیس	۰٪	۰/۲۱	۰/۴۸	۰/۳۱
	۱۰٪	۰/۲۰	۰/۶۶	۰/۳۷
	۱۵٪	۰/۲۰	۰/۵۱	۰/۵۷
حجم الیاف فولادی	۰	۰/۲۰	۰/۹۸	۰/۹۶
	۰/۴	۰/۲۱	۰/۷۷	۰/۹۶
	۰/۸	۰/۲۰	۰/۹۸	۰/۸۴
	۱/۲	۰/۲۰	۰/۹۹	۰/۹۳

بحث و نتیجه‌گیری

برای بررسی فرضیه برابری میانگین‌ها، در آزمون‌های کلاسیک و آزمون بوت‌استرپ پارامتری به فرضیه نرمال بودن توزیع مشاهدات نیاز است. نتایج نشان می‌دهند برای تعداد تیمار کم و اندازه نمونه کوچک آزمون

فیشر، پیشنهاد بهترین آزمون است. ولی با افزایش تعداد نمونه‌ها و تعداد تیمارها این آزمون نمی‌تواند اندازه آزمون را حتی در حالت برابری واریانس‌ها کنترل کند. آزمون کاکران تنها آزمونی است که برای تعداد تیمار مختلف و اندازه نمونه متفاوت و واریانس‌های مختلف در سطح $0/05$ فرضیه برابری میانگین‌ها را می‌پذیرد. طبق نتایج شبیه‌سازی آزمون کریشنامورتی و همکاران که ویرایش بوت‌استرپ آزمون کاکران است برای تعداد تیمار زیاد مانند ۶ و ۱۰ در حالت نابرابری واریانس‌ها هنگامی که پراکنندگی اندازه نمونه‌ها کم باشد، بهترین آزمون است حتی در حالت برابری واریانس‌ها یکی از پیشنهادهای بهترین آزمون از نظر ماکزیم توان است. از نتایج شبیه‌سازی مشاهده شد برای تعداد تیمار کم مانند ۳ از بین ۹ آزمون در حالت نابرابری واریانس‌ها برای حالت‌های مختلف اندازه نمونه، تنها آزمون‌های ولج و جیمز و ویراهاندی به‌عنوان بهترین آزمون شناخته می‌شوند و با افزایش تعداد تیمارها آزمون کریشنامورتی و ولج تعدیل یافته به آن‌ها اضافه می‌شود. در حالت برابری واریانس‌ها برای تعداد تیمارها و اندازه نمونه‌های مختلف، آزمون‌های ویراهاندی، ولج و ولج تعدیل یافته بهترین آزمون‌ها هستند این در حالی است که آزمون فیشر در این حالت تنها برای ۳ تیمار و اندازه نمونه کوچک و برابر، بهترین آزمون است.

منابع

1. O. Asiribo, J.Gurland, "Coping with variance heterogeneity", Computational in Statistics: Theory and Methods, 19 (1990) 4029-4048.
2. M. B. Brown, A. B. Forsythe, "The small sample behavior of some statistics which test the equality of several means", Technometrics, 16 (1974) 129-132.
3. W. G. Cochran, "Problems arising in the analysis of a series of similar experiments", Journal of Statistical Society-Supplement, 4 (1937) 102-118.
4. J. Hartung, D. Argac, K. H. Makambi, "Small sample properties of test on homogeneity in one-way ANOVA and meta-analysis. Statistical papers, 43 (2002) 197-235.
5. G. S. James, "The comparison of several groups of observation when the ratios of population variances are unknown", Biometrika, 38 (1951) 324-329.
6. K. Krishnamoorthy, F. Lu, T. Mathew, "A parametric bootstrap approach for ANOVA with unequal variances: Fixed and random models", Computational Statistics and Data Analysis, 51 (2007) 5731-5742.
7. D. V. Mehrotra, "Improving the Brown-Forsythe solution to the generalized Behrens-Fisher problem, Computational in Statistics: Simulation and Computation, 26 (1997) 1139-1145.
8. J. Park, D. Park, "Testing the equality of a large number of normal population mean", Computational Statistics and Data Analysis, 56 (2012) 1131-1149.
9. S. Weerahandi, "ANOVA under unequal error variances", Biometrics, 51 (1995) 589-599.
10. B. L. Welch, "On the comparison of several mean values: an alternative approach", Biometrika, 38 (1951) 330-336.