مجله علوم آماری، پاییز و زمستان ۱۳۹۷
جلد ۱۲، شماره ۲، ص ۳۲۳ – ۳۴۰
DOI: 10.29252/jss.12.2.323

مقایسه آزمون‌های همگنی واریانس‌ها در طرح بلوکی کامل تصادفی

نبز اسمعیل‌زاده، رضا نیک‌بخت

گروه آمار، دانشگاه کردستان

تاریخ دریافت: ۱۳۹۶/۰۳/۰۴ تاریخ آخرین بازنگری: ۱۳۹۷/۰۴/۲۹

چکیده: آزمون همگنی واریانس‌ها اغلب به عنوان یک آزمون مقدماتی برای تحلیل‌های دیگر، مثل آزمون برابری میانگین‌های جوامع مورد استفاده قرار می‌گیرد. تاکنون آزمون‌های متعددی در طرح بلوکی کامل تصادفی به منظور بهبود توان و ارائه‌ی یک آزمون استوار در برابر فرض نرمال بودن ارائه شده است، که رایج‌ترین آن‌ها آزمون بارتلت و لون بوده و بقیه تعمیمی از این دو آزمون هستند. توزیع آماره‌های این آزمون‌ها به‌صورت مجانبی محاسبه می‌شود. اخیرا آزمونی براساس برآورد مقادیر بحرانی معرفی شده است. در این مقاله با استفاده از شبیه‌سازی روش برآورد مقادیر بحرانی را روی نه آزمون همگنی واریانس‌ها اعمال کرده و عملکرد آن‌ها در توزیع‌های نرمال و تی با تعداد گروه‌های تیماری و بلوکی مختلف سنجیده می‌شود. نتایج نشان می‌دهند که آزمون‌ها براساس روش برآورد مقادیر بحرانی عملکرد بهتری از لحاظ حفظ نرخ خطای نوع اول و توان نسبت به حالت استفاده از توزیع تقریبی دارند.

واژه‌های کلیدی: آزمون لون، آزمون هان، آزمون بهاندری و دی، طرح بلوکی کامل تصادفی، همگنی واریانس‌ها.

۱ مقدمه

تحلیل واریانس بی شک یکی از شاخه‌های پرکاربرد در تحلیل آماری است. یکی از فرض‌های زیربنایی در مدل‌های خطی مورد استفاده در تحلیل واریانس، فرض نرمال بودن خطاها با واریانس‌های یکسان است. آماره‌های آزمون فراوانی در یک مدل تحلیل واریانس یک راهه برای بررسی فرض همگنی واریانس‌ها ارائه

آدرس الکترونیک مسئول مقاله: نبز اسمعیل‌زاده، n.esmailzadeh@uok.ac.ir

کد موضوع‌بندی ریاضی (۲۰۱۰): 62K10، 62F03

شده است. همچنین مطالعات شبیه‌سازی زیادی برای مقایسه این آزمون‌ها از نظر استواری در مقابل فرض غیر نرمال بودن از لحاظ حفظ نرخ خطای نوع اول و توان بالا، انجام شده است. از بین آنها می‌توان به کانوور و همکاران (۱۹۸۱)، وانگ و همکاران (۲۰۱۷) و گوکپینارا و گوکپینارا (۲۰۱۷) اشاره کرد.

برای آزمون همگنی واریانس‌ها در طرح بلوکی کامل تصادفی تعداد آزمون‌های موجود برای بررسی همگنی واریانس‌های تیماری نسبتا کمتر بوده و مطالعه مقایسه‌ای اندکی صورت گرفته است (اسچالج و دسپین، ۱۹۹۶). آزمون همگنی واریانس‌ها در طرح بلوکی کامل تصادفی فرض برابری واریانس‌های $(t \geq ۲)$ گروه تیماری را در برابر اینکه حداقل ۲ گروه از آن‌ها برابر نباشند، آزمون می‌کند. آماره‌های بیشتر آزمون‌هایی که تاکنون برای همگنی واریانس‌ها در طرح بلوکی کامل تصادفی بررسی شده‌اند، دارای یک توزیع تقریبی هستند. اخیرا بهاندری و دی (۲۰۱۳) آماره‌ای پیشنهاد کردند که آن را $F_{\max}$ نامیدند. این آزمون بر اساس برآورد مقادیر بحرانی (ECV) است. همچنین براساس یک مطالعه شبیه‌سازی شده از توزیع نرمال، آنها ادعا کردند که آزمون پیشنهادی در بعضی حالات عملکرد بهتری دارد.

در این مقاله، روش برآورد مقادیر بحرانی را روی آزمون‌های لون (۱۹۶۰)، هان (۱۹۶۹)، شوکلا (۱۹۸۲)، یتنوسومارتو (۱۹۸۶) و اونیل و متیوس (۲۰۰۲) انجام داده و با استفاده از مطالعات شبیه‌سازی شده به بررسی عملکرد آن‌ها براساس توزیع نرمال می‌پردازیم. همچنین به منظور یافتن آزمون با توان بالا و استوار در برابر انحراف از فرض نرمال بودن از نظر حفظ نرخ خطای نوع اول، مقایسه‌ها برای توزیع دم سنگین تی در حالت‌های استفاده از برآورد مقادیر بحرانی و ناحیه بحرانی بر اساس توزیع تقریبی آماره‌ها انجام می‌شود. در بخش ۲، آزمون‌های مختلف همگنی واریانس‌ها را در RCBD مرور می‌کنیم. در بخش ۳ روش برآورد مقادیر بحرانی را توضیح می‌دهیم. در بخش ۴ از طریق مطالعه‌ی شبیه‌سازی خطای نوع اول و توان، روش برآورد مقادیر بحرانی آزمون‌ها را با حالت مقدار بحرانی بر اساس توزیع تقریبی می‌سنجیم. با استفاده از یک مطالعه موردی روش برآورد ناحیه بحرانی در بخش ۵ تشریح می‌شود.

۲ آزمون‌های مورد مطالعه

در یک طرح بلوکی کامل تصادفی، فرض می‌شود t تیمار و b بلوک وجود دارد. مدل خطی مناسب با اثرات ثابت به صورت

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}, \quad i = ۱, ..., t; \; j = ۱, ..., b,$$

است، که در آن y_{ij} پاسخ متناظر iامین تیمار و jامین بلوک، τ_i و β_j به ترتیب اثرات تیمار و بلوک، ε_{ij} خطای تصادفی با توزیع $N(0,\sigma_i^2)$ است و آزمون فرضیه $H_0 : \sigma_1^2 = \ldots = \sigma_t^2$ در مقابل $H_1 : \sigma_i^2 \neq \sigma_j^2$ مورد نظر است.

آزمون لون: اولین آزمون رایج، آزمون لون (۱۹۶۰) با آماره

$$F = \frac{\frac{\sum_{i=1}^{t} b(\bar{r}_{i.}-\bar{r}_{..})^2}{t-1}}{\frac{\sum_{j=1}^{b}\sum_{i=1}^{t}(r_{ij}-\bar{r}_{i.})^2}{tb-t}}$$

دارای توزیع $F_{(t-1),(tb-t)}$ است، که در آن

$$r_{ij} = |y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}|, \quad \bar{r}_{i.} = \frac{\sum_{j=1}^{b} r_{ij}}{b}, \quad \bar{r}_{..} = \frac{\sum_{i=1}^{t} \bar{r}_{i.}}{t}.$$

در حقیقت این آزمون، انجام آزمون تحلیل واریانس روی متغیرهای تبدیل یافته r_{ij} است.

آزمون هان: هان (۱۹۶۹) آزمون‌های متعددی را برای فرض همگنی واریانس‌ها در RCBD گسترش داد که به اختصار بررسی می‌شوند. اگر R^* ضریب همبستگی چندگانه نمونه باشد آماره این آزمون که به صورت

$$F = \frac{R^{*2}}{1-R^{*2}}\frac{b-t}{t-1},$$

است، تحت فرضیه H_0 از توزیع F با $(t-1, b-1)$ درجه آزادی پیروی می‌کند. آماره آزمون bR^{*2} تحت H_0 به طور مجانبی دارای توزیع χ^2 با $t-1$ درجه آزادی است. هان (۱۹۶۹) همچنین آزمون دیگری را ارائه داد که بر مبنای توان دوم مانده‌های درون تیمار ب به صورت $F_{\max} = \max_{i,k} \frac{s_i^2}{s_k^2}$ تعریف می‌شود، که در آن

$$s_i^2 = \frac{1}{b-1}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2; \quad i = 1, ..., t. \qquad (1)$$

مقدار بحرانی آزمون $F_{\max}$ را می‌توان به صورت $\exp(\omega_\alpha\sqrt{\frac{۲}{b-۱}[۱-(t-۱)^{-۲}]})$ نوشت که در آن ω_α نقطه‌ی$۱۰۰(۱-\alpha)$ درصد توزیع دامنه نمونه‌ای با حجم t از توزیع نرمال است.

آزمون شوکلا: آماره آزمون شوکلا (۱۹۸۲) به صورت

$$q_s = \frac{(t-۱)(b-۱)[t\ln(\bar{s}^۲)-\sum_{i=۱}^{t}\ln(s_i^۲)]}{t-۱+\frac{t^۲-۱}{۳t(b-۱)}-\frac{۲(t^۴-۱)}{۱۵t^۳(b-۱)^۳}-(b-۱)A_t},$$

است، که در آن $\bar{s}^۲=\sum_{i=۱}^{t}\frac{s_i^۲}{t}$، $s_i^۲$ در (۱) تعریف شده و

$$A_t = \frac{\binom{t}{۲}}{(t-۱)^۳\{۱+\frac{t(b-۱)}{۲}\}}+\frac{۲\binom{t}{۳}}{(t-۱)^۳}[\frac{۱}{۱+\frac{t(b-۱)}{۲}}-\frac{۱}{۲+\frac{t(b-۱)}{۲}}]$$

توزیع این آماره نیز تحت فرضیه $H_۰$ به طور مجانبی به $\chi^۲_{t-۱}$ همگرا است.

آزمون یتنوسومارتو: این آزمون همان آزمون لون است که تحلیل واریانس را بر اساس مقدار مطلق مانده‌ها یعنی $|r_{ij}|$ یا توان دوم آن‌ها $r_{ij}^۲$ به کار خواهد برد. تفاوت در درجات آزادی تعدیل شده توسط یتنوسومارتو (۱۹۸۶) است. آماره آزمون F به طور تقریبی به توزیع F با درجه آزادی تیماری $\frac{bt(t-۲)}{(b-۱)(t-۱)}$ و درجه آزادی باقیمانده $\frac{b(b-۲)t(t-۲)}{(b-۱)(t-۱)}$ همگرا است.

آزمون اونیل و متیوس: اونیل و متیوس (۲۰۰۲) حالتی از کمترین توان دوم‌های موزون را برای طرح‌های بلوکی کامل بر اساس آزمون لون ارائه دادند. در طرح بلوکی کامل تصادفی، کمترین توان دوم‌های موزون می‌تواند همانند کمترین توان دوم‌های رایج آماره‌های آزمون F ضرب در یک ضریب ω باشند، که این ضریب اثر بلوکی و تیماری را در بر می‌گیرد و در جدول ۱، اونیل و متیوس (۲۰۰۲) یافت می‌شود. کمترین توان دوم موزون آزمون F بر اساس مقدار مطلق باقیمانده‌های استاندارد شده، یعنی $\frac{|y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..}|}{\sqrt{۱-\frac{۱}{b}}}$ است. آماره آزمون کمترین توان دوم موزون F در توزیع به $F_{(t-۱),(t-۱)(b-۱)}$ همگرا است.

آزمون بهاندری و دی: بهاندری و دی (۲۰۱۳) آزمون جدیدی را برای همگنی واریانس‌ها در طرح‌های بلوکی کامل ارائه کردند. مجموعه $S=\{(i_۱,i_۲):۱\leq i_۱<i_۲\leq t\}$ را در نظر بگیرید. دقت شود

که مجموعه S، دارای $r = \binom{t}{2}$ عضو است. برای هر $(i_1, i_2) \in S$ قرار داده می‌شود

$$\psi_1^{(i_1,i_2)} = \sum_{i=i_1,i_2} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2, \quad \psi_2^{-(i_1,i_2)} = \sum_{\substack{i=1 \\ i \neq i_1,i_2}}^{t} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2,$$

و آماره‌های

$$F^*_{(i_1,i_2)} = \frac{\frac{\psi_1^{(i_1,i_2)}}{(b-1)}}{\frac{\psi_2^{-(i_1,i_2)}}{(t-3)(b-1)}},$$

تعریف و با F_1^*، ...، F_r^* نشان داده می‌شوند. تحت فرضیه H_0 این آماره‌ها دارای توزیع $F_{b-1,(t-3)(b-1)}$ هستند. فرض کنید $F_1'^* = \frac{1}{F_1^*}$، $F_2'^* = \frac{1}{F_2^*}$، $\cdots$، $F_r'^* = \frac{1}{F_r^*}$ وارون‌های این r آماره باشند. آماره آزمون برای آزمون H_0 در مقابل H_1 به صورت $F_{\max} = \max\{F_1^*, ..., F_r^*, F_1'^*, ..., F_r'^*\}$ است. H_0 رد می‌شود اگر $F_{\max} > C$، که در آن C مقدار بحرانی است. بهاندری و دی (۲۰۱۳) توزیع آماره $F_{\max}$ را به‌دست نیاوردند و برای تعیین مقدار C از برآورد مقادیر بحرانی تحت توزیع نرمال استفاده کردند. آنها برای $\alpha = 0/05$ جدولی تهیه کرده‌اند. همچنین در این مقاله برای $\alpha = 0/01$ برآورد مقادیر بحرانی آماره برحسب ترکیبات مختلف تیمار و بلوک در جدول ۱ ارائه شده است.

۳ برآورد مقادیر بحرانی

تمام آماره‌های آزمون‌های مورد بررسی در بخش ۲ دارای یک توزیع تقریبی هستند که هنگام تصمیم‌گیری در مورد فرض صفر یک خطای اضافی ایجاد می‌کنند. به منظور حذف یا کاهش این خطا، درصدهای تجربی آماره‌ی آزمون مبتنی بر توزیع نرمال استاندارد می‌تواند به عنوان مقادیر بحرانی در قاعده رد کردن آزمون‌ها مورد استفاده قرار گیرد. اگر توزیع تقریبی آماره‌ی آزمون به اندازه کافی مناسب نباشد، از طریق استفاده از مقادیر بحرانی برآورد شده بهبودی به دست خواهد آمد.

جدول ۱: مقادیر بحرانی C برای آزمون بهاندری و دی و $\alpha =$ ۰/۰۱

بلوک/تیمار	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
۵	۹/۲۷	۱۰/۱۵	۱۰/۱۷	۱۰/۰۳	۹/۸۱	۹/۹۱	۹/۸۶	۱۰/۳۲	۹/۹	۱۰/۱۷	۱۰/۱۶	۱۰/۷
۶	۷/۰۶	۷/۸۳	۷/۹۵	۷/۹۶	۷/۹۸	۸/۱۱	۸/۰۱	۸/۰۶	۷/۹۹	۸/۰۹	۸/۲	۸/۱۶
۷	۵/۷۹	۶/۴۲	۶/۴۲	۶/۷۴	۶/۸۹	۶/۹۹	۷/۰۴	۷/۰۷	۷/۱۱	۷/۰۷	۷/۰۸	۷/۲۲
۸	۴/۷۶	۵/۷۹	۵/۸۴	۶/۲	۶/۲۳	۶/۳۲	۶/۲۲	۶/۴	۶/۴۳	۶/۴۷	۶/۴۹	۶/۵
۹	۴/۳۷	۵/۰۸	۵/۴۹	۵/۵۲	۵/۷۶	۵/۸۵	۵/۸۴	۵/۸۶	۶/۰۴	۶/۰۱	۶/۰۵	۶/۰۱
۱۰	۳/۹۶	۴/۸۱	۵/۰۵	۵/۲۴	۵/۳۲	۵/۴۸	۵/۴۶	۵/۴۸	۵/۵۷	۵/۶۱	۵/۶۸	۵/۷۳
۱۱	۳/۷۴	۴/۴۸	۴/۶۹	۴/۸۱	۵/۰۵	۵/۲۲	۵/۲۱	۵/۲	۵/۳۱	۵/۳۴	۵/۳۷	۵/۴۳
۱۲	۳/۴۸	۴/۲۱	۴/۴۹	۴/۷۷	۴/۷۶	۴/۸۵	۴/۹۹	۵	۵/۰۵	۵/۱	۵/۲	۵/۱۸
۱۳	۳/۲۹	۴/۰۱	۴/۳۶	۴/۵۱	۴/۵۸	۴/۷۱	۴/۸	۴/۸۱	۴/۸۵	۴/۹۱	۴/۹۵	۴/۹۸
۱۴	۳/۰۷	۳/۸۲	۴/۱۵	۴/۲۸	۴/۳۹	۴/۵۸	۴/۵۷	۴/۶	۴/۶۳	۴/۷۲	۴/۸۲	۴/۷۶
۱۵	۲/۹۹	۳/۶۸	۴/۰۱	۴/۱۹	۴/۳۶	۴/۳۵	۴/۴۶	۴/۵۱	۴/۵۲	۴/۵۷	۴/۶۵	۴/۶۶
۱۶	۲/۷۸	۳/۵۴	۳/۸۴	۳/۹۹	۴/۱۸	۴/۲۴	۴/۳۴	۴/۳۸	۴/۴۲	۴/۴۷	۴/۵۲	۴/۵
۱۷	۲/۷۳	۳/۴۵	۳/۷۶	۳/۹۳	۴/۰۶	۴/۱۵	۴/۲	۴/۳۱	۴/۲۷	۴/۳۶	۴/۳۹	۴/۳۸
۱۸	۲/۶۴	۳/۳۵	۳/۶۳	۳/۸	۳/۹۲	۴/۰۷	۴/۱۲	۴/۱۸	۴/۲۱	۴/۲۷	۴/۳۲	۴/۲۸
۱۹	۲/۶	۳/۲۵	۳/۶	۳/۷۶	۳/۸۹	۳/۹۱	۴/۰۲	۴/۰۴	۴/۱۴	۴/۱۵	۴/۱۸	۴/۲۵
۲۰	۲/۴۷	۳/۲۱	۳/۴۸	۳/۷	۳/۸۳	۳/۸۸	۳/۹۵	۳/۹۶	۴/۰۲	۴/۰۷	۴/۱۱	۴/۱۳
۲۵	۲/۲۶	۲/۸۸	۳/۱۷	۳/۳۷	۳/۴۷	۳/۵۷	۳/۶۳	۳/۶۹	۳/۷۴	۳/۷۸	۳/۸	۳/۷۹
۳۰	۲/۱	۲/۷۱	۲/۹۷	۳/۲	۳/۲۷	۳/۳۵	۳/۴۳	۳/۴۹	۳/۵۱	۳/۵۲	۳/۵۷	۳/۵۹
۳۵	۱/۹۹	۲/۵۵	۲/۸۳	۳	۳/۰۹	۳/۱۹	۳/۲۴	۳/۳	۳/۳۵	۳/۳۸	۳/۴۳	۳/۴۲
۴۰	۱/۹۲	۲/۴۳	۲/۷۲	۲/۸۹	۲/۹۸	۳/۰۶	۳/۱۳	۳/۱۷	۳/۱۹	۳/۲۵	۳/۲۶	۳/۲۹
۴۵	۱/۸۲	۲/۳۵	۲/۶۲	۲/۷۹	۲/۹	۲/۹۷	۳/۰۳	۳/۰۸	۳/۱۱	۳/۱۵	۳/۱۷	۳/۲
۵۰	۱/۷۵	۲/۲۸	۲/۵۴	۲/۷	۲/۸۲	۲/۸۸	۲/۹۴	۲/۹۹	۳/۰۲	۳/۰۶	۳/۰۷	۳/۱۱
۱۰۰	۳/۱۱	۱/۹۶	۲/۱۷	۲/۳۳	۲/۴۳	۲/۴۸	۲/۵۳	۲/۵۸	۲/۶۳	۲/۶۴	۲/۶۷	۲/۶۹
۲۰۰	۱/۳۲	۱/۷۵	۱/۹۶	۲/۰۸	۲/۱۷	۲/۲۴	۲/۲۹	۲/۳۲	۲/۳۵	۲/۳۷	۲/۳۹	۲/۴۲

مراحل کار عبارتند از:

۱. برای n_i ، $i=۱,\dots,k$، نمونه‌ها از یک جامعه نرمال استاندارد تولید شود.

۲. مقدار آماره T محاسبه شود.

۳. این فرایند $۱۰۰M$ بار تکرار شود، که در آن M یک عدد صحیح است.

۴. مقادیر آماره T مرتب و با $T(۱),\dots,T(۱۰۰M)$ نشان داده شوند.

۵. مقدار بحرانی ۱۰۰α درصد به صورت $C=\frac{۱}{۲}[T((۱-۱۰۰\alpha)M)+T((۱-۱۰۰\alpha)M+۱)]$ محاسبه شود. اگر مقدارآماره آزمون بزرگتر از C باشد فرضیه صفر رد شود.

برای تولید درصدهای تجربی در تمام موارد از توزیع نرمال استاندارد استفاده می‌شود که ممکن است خطایی را تولید کند که به میزان حساسیت آماره آزمون به توزیع‌های غیرنرمال بستگی دارد. با این حال، زوقتی که توزیع واقعی داده‌ها مشخص است، آزمون را می‌توان با استفاده از مقادیر بحرانی برآورد شده با استفاده از توزیع اصلی به صورت دقیق‌تر انجام داد.

۴ مطالعات شبیه‌سازی برای مقایسه آزمون‌ها

برای بررسی‌های شبیه‌سازی تحت فرض برابری واریانس‌ها، حالت‌هایی انتخاب می‌شوند که شامل ۵ و ۱۰ گروه تیماری باشد، تعداد بلوک‌های متناظر با گروه‌های تیماری ۵، ۱۰ و ۲۰ است. برای دستیابی به آزمون‌های استوار و پرتوان نمونه‌ها از چهار توزیع با ویژگی‌های مختلف تولید شده و مورد بررسی قرار می‌گیرد. برای این منظور از شبیه‌سازی مونت کارلو برای بررسی نرخ خطای نوع اول و توان آزمون‌ها بهره برده خواهد شد. برای محاسبه نرخ‌های رد شبیه‌سازی شده تحت فرض $H_۰$، نرخ رد فرض صفر را ۰٫۰۵ قرار داده و این فرایند ۱۰۰۰۰ بار تکرار می‌شود. ترکیبات مختلف تیماری و بلوکی با توزیع‌های نرمال (متقارن)، تی استیودنت (متقارن، دنباله‌ی بلند وکشیدگی کوتاه)، در نظر گرفته شده است.

نتایج شبیه‌سازی به اختصار در جدول ۲ و شکل‌های ۱ تا ۴ آورده شده‌اند. پنج قسمت جدول به ترتیب متناظر با ترکیبات (بلوک، تیمار)=(۵،۱۰)، (۵،۲۰)، (۱۰،۵)، (۱۰،۱۰) و (۱۰،۲۰) هستند. تحت فرض صفر واریانس‌ها برابر یک قرار داده شده‌اند. در هر قسمت جدول ۲، انحراف معیارها تحت فرض مقابل برای تیمار ۱ به ترتیب برابر ۱٫۵، ۲، ۲٫۵ و ۳ و برای بقیه تیمارها برابر ۱ در نظر گرفته شده‌اند. در هر

قسمت، سطر اول شبیه‌سازی خطای نوع اول و چهار سطر بعدی توان آزمون‌ها را نشان می‌دهند. لازم به ذکر است وقتی تعداد گروه‌های تیماری بزرگ‌تر از تعداد بلوک‌ها باشد دو آزمون دقیق و مجانبی هان قابل محاسبه نیستند. طبق معیار برادلی آزمونی را استوار گوییم هرگاه برای سطح اسمی $\alpha = ۰/۰۵$، نرخ خطای نوع اول آن در فاصله [۰/۰۲۵, ۰/۰۷۵] قرار گیرد.

۱.۴ نتایج شبیه سازی نرخ خطای نوع اول

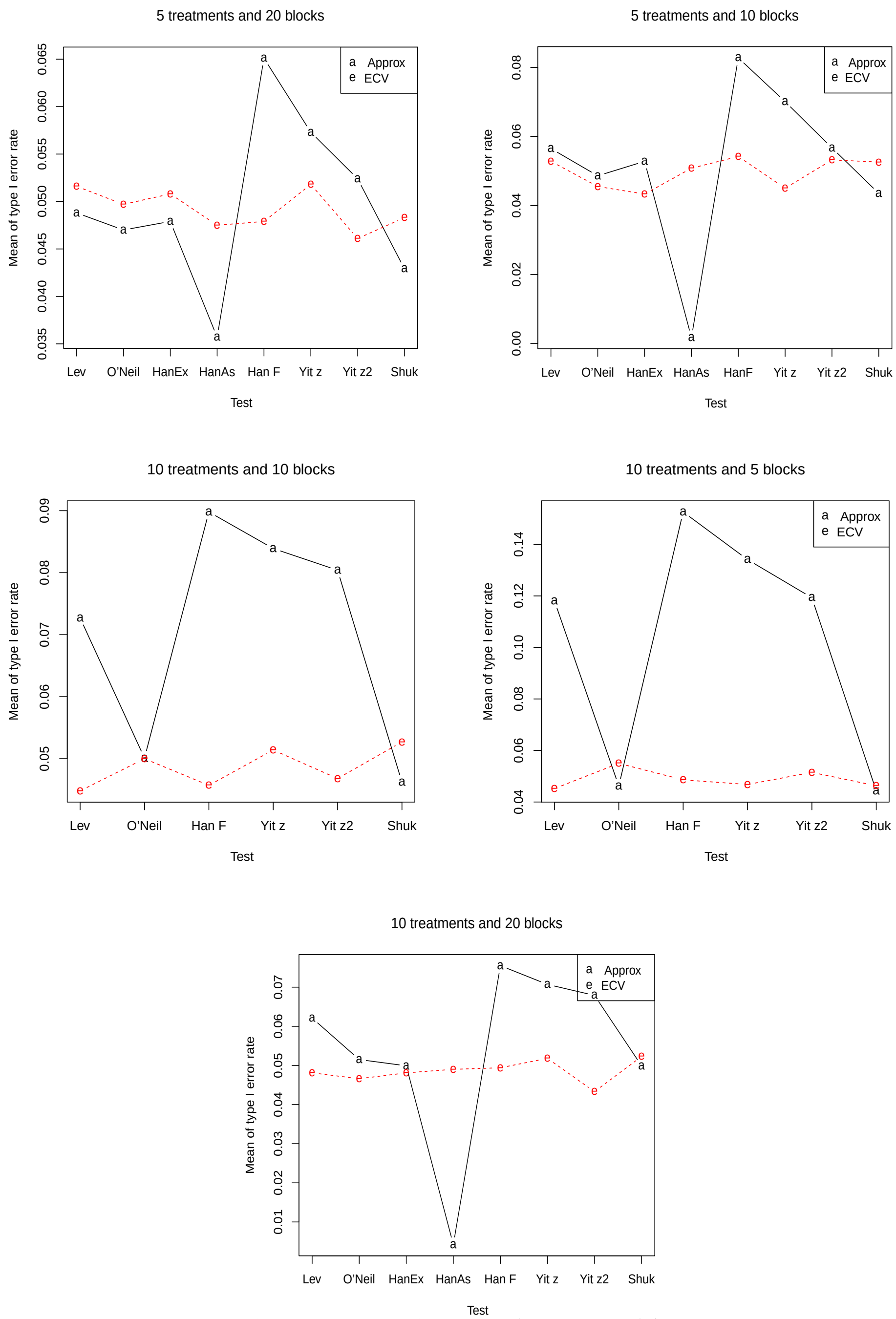
در این بخش نتایج شبیه‌سازی نرخ خطای نوع اول براساس ترکیبات مختلف تیمار و بلوک و توزیع‌های گوناگون بررسی می‌شود.

۱.۱.۴ حالت پنج گروه تیماری و ده بلوک

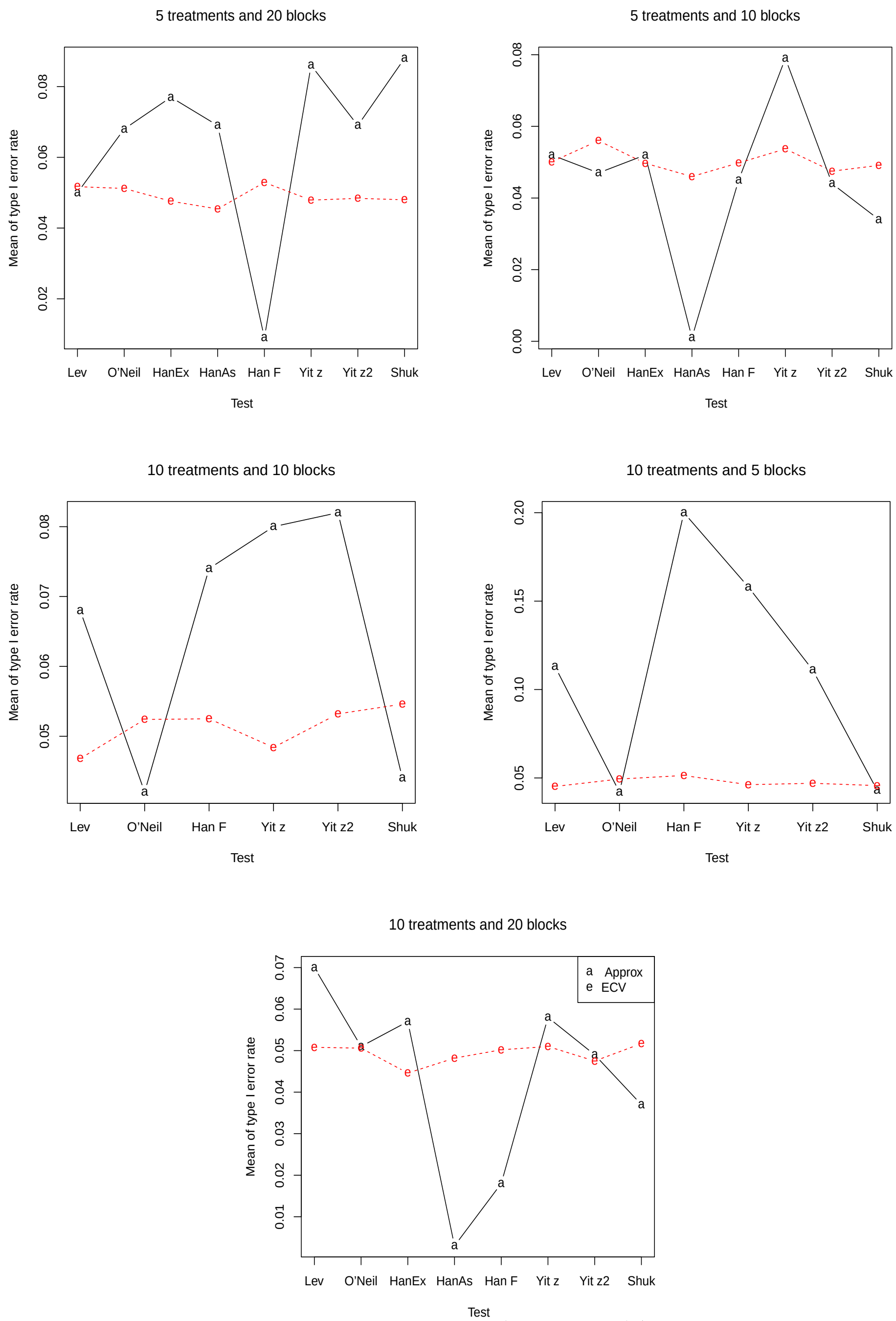
توزیع نرمال: در این حالت نتایج در جدول ۲ و شکل ۱ نشان می‌دهند که تحت روش ناحیه بحرانی بر اساس توزیع تقریبی آزمون‌های $F_{\max}$ هان و z یتنوسومارتو قادر نیستند که خطای نوع اول را به مقدار $\alpha = ۰/۰۵$ نزدیک نمایند. روش برآورد مقادیر بحرانی برای تمامی آزمون‌ها به طور کامل استوار بوده و نسبت به روش ناحیه بحرانی بر اساس توزیع تقریبی برتری قابل ملاحظه‌ای دارد.
توزیع تی استیودنت: در این حالت روش ناحیه بحرانی بر اساس توزیع تقریبی عملکرد مناسبی دارد، با این وجود می‌توان دید که آزمون z یتنوسومارتو نرخ خطای بالایی دارد و استوار نیست. اما در روش برآورد مقادیر بحرانی مشاهده می‌کنیم که مقادیر نرخ خطا به مقدار $\alpha = ۰/۰۵$ نزدیک هستند. از این رو تمامی آزمون‌ها با این روش کاملا استوار بوده و کنترل خوبی روی نرخ خطای نوع اول دارند (شکل ۲).

۲.۱.۴ حالت پنج گروه تیماری و بیست بلوک

توزیع نرمال: با توجه به شکل ۱، نرخ خطای نوع اول در حالت توزیع تقریبی تا حدی کنترل شده است، هر چند آزمون مجانبی و $F_{\max}$ هان وضعیت مناسبی ندارند. روش برآورد مقادیر بحرانی در این حالت نیز وضعیت مطلوبی دارد به طوری که نرخ خطای نوع اول تمام آزمون‌ها نزدیک به مقدار $\alpha = ۰/۰۵$ است.
توزیع تی استیودنت: نتایج شبیه‌سازی در این حالت در شکل ۲ نشان می‌دهند که آزمون $F_{\max}$ هان در این حالت آزمونی استوار نیست، هر چند که آزمون‌های لون و اونیل نیز با این روش وضعیت چندان خوبی ندارند. روش برآورد مقادیر بحرانی برای همه‌ی آزمون‌ها نرخ خطای نوع اول را کنترل کرده است.



شکل ۱: نرخ خطای نوع اول برای توزیع نرمال



شکل ۲: نرخ خطای نوع اول برای توزیع تی

۳.۱.۴ حالت ده گروه تیماری و پنج بلوک

در این قسمت نتایج شبیه‌سازی خطای نوع اول را برای ترکیب ۱۰ گروه تیماری و ۵ بلوک ارایه می‌کنیم. در این حالت آماره‌های آزمون برای آزمون‌های دقیق و مجانبی هان قابل محاسبه نیست.

توزیع نرمال: همان طور که در جدول ۲ و شکل ۱ ملاحظه می‌شود تحت روش ناحیه بحرانی بر اساس توزیع تقریبی، نرخ خطا برای آزمون‌های لون، $F_{\max}$ هان و آزمون‌های یتنوسومارتو از ۰/۱ بیشتر است و طبق معیار برادلی استوار نیستند. ولی وقتی از روش برآورد مقادیر بحرانی استفاده می‌کنیم. بهبود عملکرد این آزمون‌ها را با توجه به نمودار می‌توان مشاهده کرد.

توزیع تی استیودنت: در این حالت عملکرد روش توزیع تقریبی تا حدی بهتر می‌شود. با این وجود، باز هم آزمون‌های یتنوسومارتو، $F_{\max}$ هان و همچنین آزمون لون با اینکه نسبت به انحراف از فرض نرمال بودن حساسیت کمتری دارد، ولی استوار نیستند. اما روش برآورد مقادیر بحرانی با کنترل نرخ خطای نوع اول آزمون‌ها، برتری خود را اثبات می‌کند (شکل ۲).

۴.۱.۴ حالت ده گروه تیماری و ده بلوک

توزیع نرمال با توجه به جدول ۲ و شکل ۱ در می‌یابیم که تحت روش ناحیه بحرانی بر اساس توزیع تقریبی آزمون‌های یتنوسومارتو و $F_{\max}$ هان استوار نبوده و نرخ رد بیشتر از ۰/۰۸ دارند. اما آزمون‌های لون، اونیل و شوکلا توانسته‌اند که نرخ خطای نوع اول را کنترل نمایند. با اعمال روش برآورد مقادیر بحرانی عملکرد مناسب آزمون‌ها مشاهده می‌شود. نرخ خطا به طور قابل ملاحظه‌ای نسبت به حالت توزیع تقریبی پایین می‌آید و برای تمامی آزمون‌ها استوار شده است.

توزیع تی استیودنت: نتایج در این حالت نشان می‌دهد که روش ناحیه بحرانی بر اساس توزیع تقریبی بهبود قابل ملاحظه‌ای پیدا می‌کند. با این حال آزمون‌های یتنوسومارتو تحت این روش استوار نیستند و آزمون $F_{\max}$ هان نیز وضعیت مناسبی ندارد. اما با اعمال روش برآورد مقادیر بحرانی نرخ خطای نوع اول برای تمامی آزمون‌ها به خوبی کنترل شده است (شکل ۲).

۵.۱.۴ حالت ده گروه تیماری و بیست بلوک

توزیع نرمال: براساس روش تقریبی هیچ‌کدام از آزمون‌های لون، یتنوسومارتو و همچنین آزمون‌های $F_{\max}$ و مجانبی هان قادر نیستند که نرخ خطای نوع اول را به خوبی کنترل کنند. با استناد به جدول ۲ و شکل ۱ می‌توان گفت که بر اساس روش برآورد مقادیر بحرانی تمامی آزمون‌ها به طور کامل استوار هستند.

جدول ۲: نرخ خطای نوع اول و توان آزمون‌ها بر اساس روش برآورد مقادیر بحرانی در توزیع نرمال

Bhand	Lev	O'Neil	HanEx	HanAs	HanF	Yit.z	Yit.z2	Shuk
۰٫۰۴۶۵	۰٫۰۵۲۸	۰٫۰۴۵۵	۰٫۰۴۳۳	۰٫۰۵۰۸	۰٫۰۵۴۲	۰٫۰۴۵	۰٫۰۵۳۲	۰٫۰۵۲۶
۰٫۱۴۵۸	۰٫۱۲۷۳	۰٫۱۲۰۱	۰٫۰۶۵۷	۰٫۰۷۷۳	۰٫۱۳۰۸	۰٫۱۲۰۲	۰٫۱۳۴۷	۰٫۱۴۶۳
۰٫۳۹۵۳	۰٫۳۲۸۵	۰٫۳۲۹۷	۰٫۱۳۰۹	۰٫۱۳۱۱	۰٫۳۲۹۹	۰٫۳۲	۰٫۳۴۴۷	۰٫۳۹۶۳
۰٫۶۶۷	۰٫۵۴۰۵	۰٫۵۵۲۷	۰٫۲۱۸۴	۰٫۲۳۳۳	۰٫۵۶۸۳	۰٫۵۵۳۷	۰٫۵۷۸۶	۰٫۶۷۳۳
۰٫۸۲۹۲	۰٫۷۱۳۹	۰٫۷۲۵۴	۰٫۳۱۷۵	۰٫۳۴۳۴	۰٫۷۵۲۹	۰٫۷۲۲۸	۰٫۷۴۸۴	۰٫۸۳۹
۰٫۰۵۰۴	۰٫۰۵۱۶	۰٫۰۴۹۷	۰٫۰۵۰۸	۰٫۰۴۷۵	۰٫۰۴۷۹	۰٫۰۵۱۸	۰٫۰۴۶۱	۰٫۰۴۸۳
۰٫۲۹۵۶	۰٫۲۴۳۱	۰٫۲۴۵۲	۰٫۱۴۶۸	۰٫۱۳۲۲	۰٫۲۴۹۸	۰٫۲۴۹	۰٫۲۴۱۴	۰٫۲۸۵
۰٫۷۴۲۴	۰٫۶۶۱۱	۰٫۶۷۹۶	۰٫۴۰۳۸	۰٫۳۹۷۲	۰٫۶۸۲۱	۰٫۶۸۶	۰٫۶۶۸۶	۰٫۷۴۲۳
۰٫۹۴۹۴	۰٫۸۹۸۲	۰٫۹۰۹۴	۰٫۶۸۷۶	۰٫۶۶۵۹	۰٫۹۲۴۲	۰٫۹۱۳	۰٫۹۰۵۸	۰٫۹۴۹۱
۰٫۹۹۰۴	۰٫۹۷۵۴	۰٫۹۷۶۸	۰٫۸۶۸۳	۰٫۸۵۸۱	۰٫۹۸۶۱	۰٫۹۷۹۹	۰٫۹۷۸۲	۰٫۹۹۲۲
۰٫۰۵۵۲	۰٫۰۴۵۲	۰٫۰۵۵	NA	NA	۰٫۰۴۸۶	۰٫۰۴۶۸	۰٫۰۵۱۵	۰٫۰۴۶۳
۰٫۱۲۱۵	۰٫۰۸۰۱	۰٫۰۹۰۲	NA	NA	۰٫۰۷۴۷	۰٫۰۸۲۹	۰٫۰۹۴	۰٫۰۷۹۳
۰٫۲۹۰۸	۰٫۱۷۱۶	۰٫۱۸۸۸	NA	NA	۰٫۱۲۹۷	۰٫۱۷۸۵	۰٫۱۸۱۹	۰٫۲۰۲۲
۰٫۵۰۳۳	۰٫۳۰۳۴	۰٫۳۱۳	NA	NA	۰٫۲۲۲۳	۰٫۳۰۳۳	۰٫۳۱۸۹	۰٫۳۶۶۵
۰٫۶۵۴۷	۰٫۴۱۸۵	۰٫۴۵۳	NA	NA	۰٫۳۲۹۴	۰٫۴۲۵۷	۰٫۴۴۶	۰٫۵۲۵۴
۰٫۰۵۵۸	۰٫۰۴۴۸	۰٫۰۵	NA	NA	۰٫۰۴۵۷	۰٫۰۵۱۴	۰٫۰۴۶۸	۰٫۰۵۲۷
۰٫۲۱۶۷	۰٫۱۳۶۷	۰٫۱۴۵۳	NA	NA	۰٫۱۱۷۶	۰٫۱۴۶۴	۰٫۱۵۰۴	۰٫۱۵۷۶
۰٫۵۸۵۹	۰٫۴۰۲۶	۰٫۴۱۵۴	NA	NA	۰٫۳۶۰۵	۰٫۴۰۳۹	۰٫۴۱۳	۰٫۴۶۳۸
۰٫۸۳۹۸	۰٫۶۶۲۶	۰٫۶۷۶۹	NA	NA	۰٫۶۳۱۴	۰٫۶۶۵۶	۰٫۶۷۴۹	۰٫۷۴۲۴
۰٫۹۳۸۵	۰٫۸۲۴	۰٫۸۳۲۶	NA	NA	۰٫۸۲۱۶	۰٫۸۲۷۲	۰٫۸۳۰۴	۰٫۸۹۳
۰٫۰۴۹۸	۰٫۰۴۸۱	۰٫۰۴۶۶	۰٫۰۴۸۱	۰٫۰۴۹	۰٫۰۴۹۴	۰٫۰۵۱۸	۰٫۰۴۳۳	۰٫۰۵۲۳
۰٫۳۹۷۴	۰٫۲۹۳۴	۰٫۲۷۹۹	۰٫۰۷۱	۰٫۰۶۷۵	۰٫۲۸۶۹	۰٫۲۹۵	۰٫۲۶۶۳	۰٫۳۲۱۵
۰٫۸۸۲۴	۰٫۷۵۸۲	۰٫۷۵۸۷	۰٫۱۲۲	۰٫۱۲۲۲	۰٫۷۸۰۳	۰٫۷۷۳۱	۰٫۷۴۸۱	۰٫۸۱۳۷
۰٫۹۸۵۵	۰٫۹۵۲۴	۰٫۹۵۲۴	۰٫۲۱۱۴	۰٫۲۲۸۲	۰٫۹۶۴	۰٫۹۵۴۹	۰٫۹۵۲	۰٫۹۶۹۷
۰٫۹۹۸۵	۰٫۹۹۰۷	۰٫۹۹۱۹	۰٫۳۳۵۲	۰٫۳۵۱۵	۰٫۹۹۵۵	۰٫۹۹۰۷	۰٫۹۹۱۱	۰٫۹۹۶۵

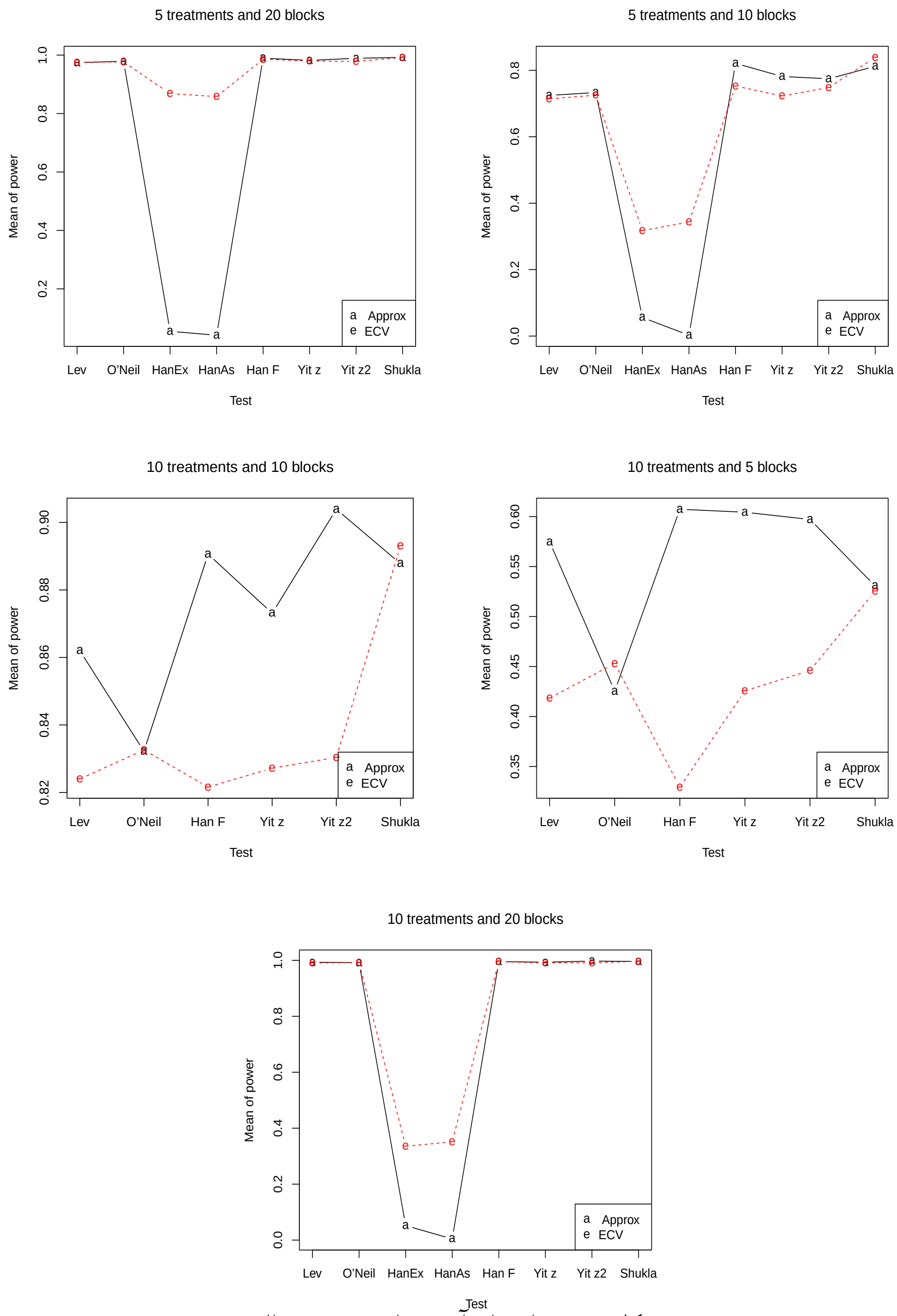
سطر اول هر قسمت جدول مربوط به خطای نوع اول و بقیه توان آزمون است .

توزیع تی استیودنت: در این حالت روش تقریبی بهبود قابل توجهی داشته است. با این وجود آزمون‌های $F_{\max}$ و مجانبی هان طبق معیار برادلی استوار نیستند و آزمون لون هم با نرخ رد فرضیه صفر ۰/۰۷ قادر نیست که خطای نوع اول را تا حدی به مقدار $\alpha = ۰/۰۵$ نزدیک کند و این مطلب ضعف این آزمون را در حالت انحراف از فرض نرمال بودن نشان می‌دهد. روش برآورد مقادیر بحرانی توانسته است برای تمامی آزمون‌ها نرخ خطای نوع اول را کنترل کند (شکل ۲).

۲.۴ مقایسه توان آزمون‌ها

در این بخش توان آزمون‌ها بر اساس توزیع تقریبی و برآورد مقادیر بحرانی بررسی می‌شود.
توزیع نرمال: مقایسه توان آزمون‌ها در حالت توزیع تقریبی و روش برآورد مقادیر بحرانی در چند بخش مورد بحث قرار می‌گیرد. در توزیع نرمال ابتدا حالتی را در نظر می‌گیریم که تعداد گروه‌های تیماری ۵ وتعداد بلوک‌ها بین ۱۰ و ۲۰ متغیر است. در این حالت زمانی که از روش برآورد مقادیر بحرانی استفاده می‌کنیم، توان آزمون‌های دقیق و مجانبی هان و آزمون شوکلا به طور چشم‌گیری بهبود می‌یابد و در سایر آزمون‌ها با روش توزیع تقریبی قابل قیاس است، این امر به وضوح در جدول ۲ و شکل ۳ مشهود است. در حالتی که تعداد گروه‌های تیماری بزرگ‌تر از تعداد بلوک‌ها است، آزمون $F_{\max}$ از عملکرد بهتری نسبت به سایر آزمون‌ها برخوردار است. در روش برآورد مقادیر بحرانی زمانی که تعداد گروه‌های تیماری و تعداد بلوک‌ها برابر هستند، آزمون‌های اونیل و شوکلا تمایل دارند که توان بیشتری نسبت به حالت توزیع تقریبی داشته باشند. زمانی که تعداد گروه‌های تیماری ۱۰ و تعداد بلوک‌ها ۲۰ است، آزمون‌های دقیق و مجانبی هان و آزمون شوکلا در روش برآورد مقادیر بحرانی توان بیشتری دارند.

توزیع تی استیودنت: نتایج در این حالت در شکل ۴ نمایش داده شده است. زمانی که تعداد گروه‌های تیماری کوچک‌تر از تعداد بلوک‌هاست، مشاهده می‌کنیم که با اعمال روش برآورد مقادیر بحرانی، توان آزمون‌های مجانبی و $F_{\max}$ هان و آزمون شوکلا به طور چشم‌گیری افزایش پیدا می‌کند و در بقیه آزمون‌ها با روش ناحیه بحرانی بر اساس توزیع تقریبی قابل قیاس است. وقتی که تعداد گروه‌های تیماری ۱۰ و تعداد بلوک‌ها ۵ است، توان آزمون‌های لون، $F_{\max}$ هان و یتنوسومارتو پایین می‌آید. زمانی که تعداد گروه‌های تیماری و بلوکی برابر هستند آزمون $F_{\max}$ هان در حالت برآورد مقادیر بحرانی توان بیشتری دارد و در بقیه آزمون‌ها می‌توان گفت که دو روش تفاوت چندانی ندارند. وقتی که تعداد بلوک‌ها به ۲۰ افزایش پیدا می‌کند، آزمون مجانبی و $F_{\max}$ هان عملکرد بهتری در حالت روش برآورد مقادیر بحرانی دارند. در این توزیع در بین آزمون‌های استوار تحت روش توزیع تقریبی، آزمون شوکلا و در بین آزمون‌هایی که در بعضی



شکل ۳: نمودار توان آزمون‌ها در توزیع نرمال

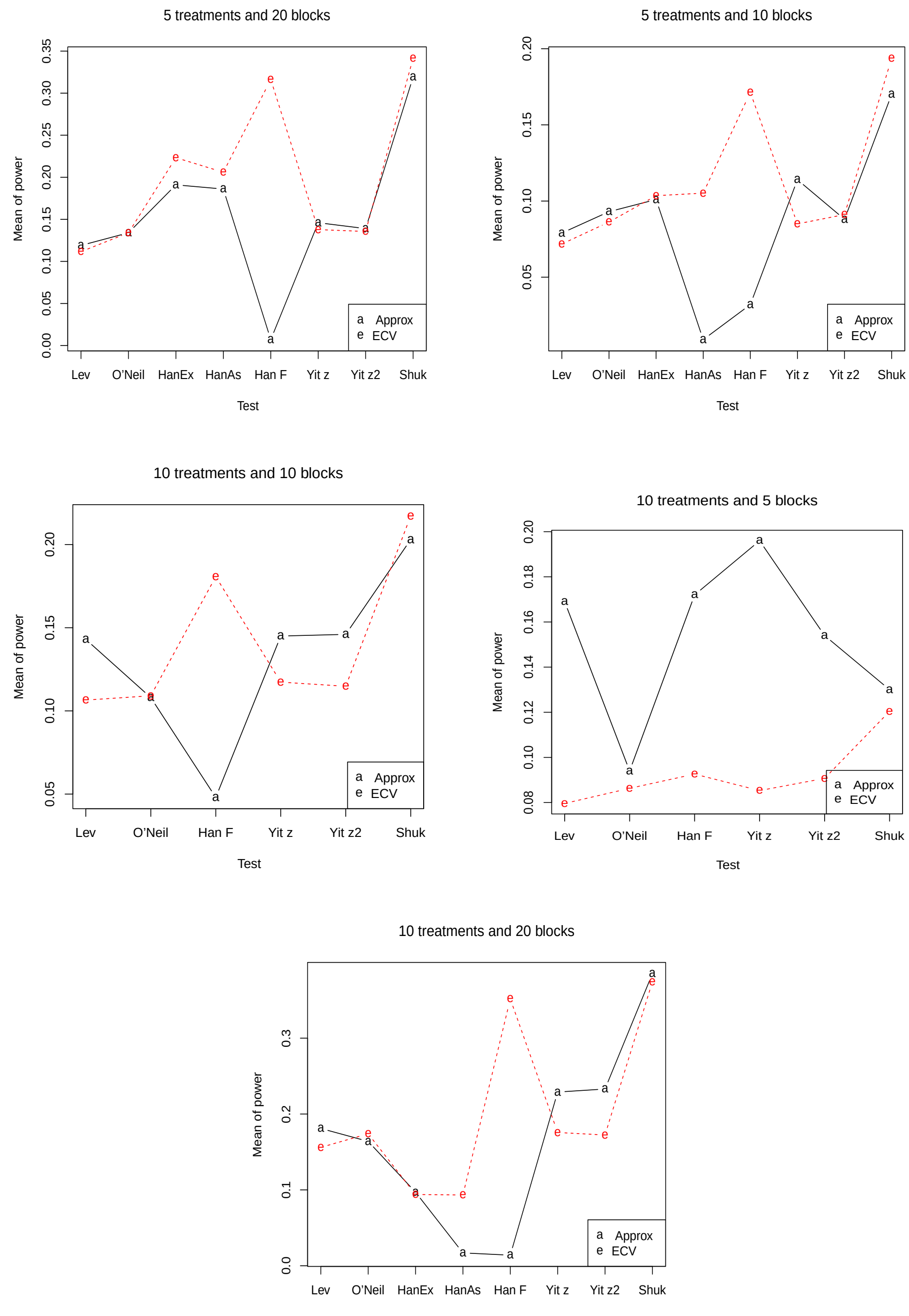
حالات ترکیبات تیماری و بلوکی غیر استوار هسند، آزمون z^2 یتنوسومارتو پر توان‌ترین آزمون است. در حالت روش برآورد مقادیر بحرانی آزمون بهاندری و دی دارای بیشینه توان است.

۵ مطالعه موردی

در این بخش با یک مثال کاربردی استفاده از روش برآورد ناحیه بحرانی برای آزمون‌های ذکر شده در بخش ۳ را تشریح می‌کنیم. آزمایشی را در نظر بگیرید که در آن اثر چهار گونه گندم که در ۱۳ ناحیه متفاوت کاشت می‌شوند، روی میزان محصول برداشتی مورد نظر است. داده‌ها در جدول ۳ آورده شده‌اند. این داده‌ها ابتدا در گریبیل (۱۹۵۴) ذکر شدند و توسط هان (۱۹۶۹)، بهاندری و دی (۲۰۱۳) و محققان زیادی تحلیل شده‌اند. بررسی نمودار احتمال نرمال مانده‌ها فرضیه نرمال بودن را تایید می‌کند. مقادیر آماره‌های آزمون مختلف و برآورد مقادیر بحرانی متناظر در جدول ۴ داده شده‌اند. روش مقادیر بحرانی در صورتی که مقدار مشاهده شده آماره از مقدار برآورد بحرانی بزرگتر باشد، فرضیه صفر را رد می‌کند. پس در این مثال همه آزمون‌ها فرضیه همگنی واریانس‌های تیماری را رد می‌کنند. این سازگار با نتایج تحقیقات قبلی است.

جدول ۳: میزان محصول برای چهار گونه گندم در ۱۳ ناحیه

ناحیه	گونه‌های گندم			
	۴	۳	۲	۱
۱	۴۳٫۶۰	۲۴٫۰۵	۱۹٫۴۷	۱۹٫۴۱
۲	۴۰٫۴۰	۲۱٫۷۶	۱۶٫۶۱	۲۳٫۸۴
۳	۱۸٫۰۸	۱۴٫۱۹	۱۶٫۶۹	۱۶٫۰۸
۴	۱۹٫۵۷	۱۸٫۶۱	۱۷٫۷۸	۱۸٫۲۹
۵	۴۵٫۲۰	۲۹٫۳۳	۲۰٫۱۹	۳۰٫۰۸
۶	۲۵٫۸۷	۲۵٫۶۰	۲۳٫۳۱	۲۷٫۰۴
۷	۵۵٫۲۰	۳۸٫۷۷	۲۱٫۱۵	۳۹٫۹۵
۸	۵۵٫۳۲	۳۴٫۱۹	۱۸٫۵۶	۲۵٫۱۲
۹	۱۹٫۷۹	۲۱٫۶۵	۲۳٫۳۱	۲۲٫۴۵
۱۰	۴۶٫۲۴	۳۱٫۵۲	۲۲٫۴۸	۲۹٫۲۸
۱۱	۱۴٫۸۸	۱۵٫۶۸	۱۹٫۷۹	۲۲٫۵۶
۱۲	۷٫۵۲	۴٫۶۹	۲۰٫۵۳	۲۲٫۰۸
۱۳	۴۱٫۱۷	۳۲٫۵۹	۲۹٫۲۵	۴۳٫۹۵



شکل ۴: نمودار توان آزمون‌ها در توزیع تی

بحث و نتیجه‌گیری

فرضیه همگنی واریانس‌ها یکی از مفروضات اصلی در تحلیل یک طرح بلوکی کامل است. آماره‌های آزمون مختلفی برای بررسی این فرضیه ارایه شده است. در این مقاله آزمون‌های لون، هان، شوکلا، یتنوسومارتو، اونیل و متیوس و بیهاندری و دی مرور شده‌اند. مقایسه عملکرد آنها از نظر کنترل نرخ خطای نوع اول و توان در شرایط گوناگون به صورت نظری و تحلیلی امکان پذیر نیست. لذا، براساس شبیه سازی مونت کارلو عملکرد آنها برای ترکیبات تیماری مختلف بلوک و تیمار بررسی می‌شود. آماره‌های آزمون مورد بررسی دارای توزیع تقریبی هستند. در این مقاله، آزمون فرض همگنی واریانس‌ها در دو حالت استفاده از ناحیه بحرانی براساس توزیع تقریبی و روش برآورد ناحیه بحرانی مورد مقایسه قرار گرفته‌اند. به منظور سنجش استواری آزمون‌ها در برابر نقض فرض نرمال بودن خطاها، خطای نوع اول و توان آماره‌ها برای توزیع تی نیز بررسی شدند. خطای نوع اول تحت روش برآورد مقادیر بحرانی به خوبی کنترل شده است. همچنین توان آزمون‌ها در این حالت یا بهتر است یا در بعضی شرایط با روش توزیع تقریبی قابل قیاس است. زمانی که تعداد گروه‌های تیماری بزرگ و تعداد بلوک‌ها به نسبت کوچک است، آزمون $F_{\max}$ که بر اساس روش برآورد مقادیر بحرانی است، از عملکرد بهتری نسبت به سایر روش‌ها برخوردار است.

جدول ۴: مقادیر آماره‌های آزمون برای داده‌های میزان محصول

آزمون	Shuk	Yit.z2	Yit.z	HanF	HanAs	HanEx	ONeil	Lev	Bhand
آماره	۱۷/۷۰	۱۳/۲۴	۲۱/۰۷	۱۲/۴۳	۱۱/۵۷	۲۴/۴۰	۱۹/۴۹	۱۲/۰۵	۴/۵۹
ECV	۷/۴۴	۲/۹۸	۳/۱۳	۴/۴۵	۷/۳۷	۳/۷۹	۲/۷۱	۲/۶۳	۲/۵۶

تقدیر و تشکر

از داوران محترم به خاطر پیشنهادات مفید در راستای بهبود متن مقاله و همچنین ویراستار محترم به خاطر دقت فراوان سپاسگزاریم.

مراجع

Bhandary, M. and Dai, H. (2013), An Alternative Test for the Equality of Variances for Several Populations in Randomized Complete Block Design, *Statistical Methodology*, **11**, 22-35.

Conover, W. J., Johnson, M. E. and Johnson, M. M. (1981), A Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer Continental Shelf Bidding Data, *Technometrics*, **23**, 351–361.

Gokpinara, E. and Gokpinara, F. (2017), Testing Equality of Variances for Several Normal Populations, *Communications in Statistics - Simulation and Computation*, **46**, 38-52.

Graybill, F. (1954), Variance Heterogeneity in a Randomized Block Design, *Biometrics*, **10**, 516–520.

Schaalje, B. G. and Despain, D. (1996), Robustness of Homogeneity of Variance Tests for Randomized Complete Block Data, *Communications in Statistics- Simulation and Computation*, **25**, 961–977.

Han, C. P. (1969), Testing the Homogeneity of Variances in a Two-Way Classification, *Biometrics*, **25**, 153-158.

Levene, H. (1960), In: Olkin I., et al. (Eds.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotellings*, Stanford University Press, 278-292.

O'Neill, M. and Mathews, K. (2002), Levene Tests of Homogeneity of Variance for General Block and Treatment Designs, *Biometrics*, **58**, 216-224.

Shukla, G. K. (1982), Testing the Homogeneity of Variances in a Two-Way Classification, *Biometrik*, **69**, 411-416.

Wang, Y., Rodriguez, P., Chen, Y., Kromrey, J. D., Kim, E. S., Pham, T., Nguyen, D. and Romano, J. L. (2017), Comparing the Performance of Approaches for Testing the Homogeneity of Variance Assumption in One-Factor ANOVA Models, *Educational and Psychological Measurement*, **77**, 305–329.

Yitnosumarto, S. (1986), Levene's Tests of Variance Homogeneity, *Australian Journal of Statistics*, **28**, 230-241.