آشنایی با روش مونت کارلو در آزمون فرض‌ها

حمزه ترابی [1]

چکیده:

روش مونت‌کارلو برای آزمون فرض‌ها بر پایه‌ی شبیه‌سازی استوار است و درآن، برخلاف روش‌های رایج کلاسیک آزمون فرض، نیازی به دانستن توزیع دقیق آماره‌ی آزمون تحت فرض صفر نداریم، ولی باید بتوانیم آماره‌ی آزمون را تحت فرض صفر شبیه‌سازی کنیم.

در این مقاله، در آغاز به معرفی روش مونت‌کارلو در آزمون فرض‌ها می‌پردازیم و سپس برای مثال این روش را برای دو آزمون فرض به‌کار می‌بریم. برای مقایسه‌ی روش مونت‌کارلو و روش‌های کلاسیک آزمون فرض، یک مسئله‌ی آزمون فرض را با بهره از دو روش حل می‌نماییم.

واژه‌های کلیدی: شبیه‌سازی ایستا و پویا، آزمون فرض، روش مونت‌کارلو.

۱ مقدمه

یک مسئله‌ی آزمون فرض، بررسی ادعا، گمان و حدسی درباره توزیع یا پارامترهای توزیع جامعه است. فرض کنید می‌خواهیم فرض صفر H_0 را در برابر H_1 بیازماییم و آماره‌ی آزمون نیز U و مقدار مشاهده شده‌ی آن از نمونه‌ی تصادفی، u_1 باشد. روشن است که اگر توزیع آماره‌ی U تحت فرض H_0 معلوم باشد، می‌توان پس از تعیین ناحیه‌ی بحرانی، آزمونی با اندازه‌ی آزمون (یا سطح آزمون) دلخواه انجام داد. ولی در بیشتر مسائل آزمون فرض، با آماره‌های آزمونی سر و کار داریم که یافتن توزیع آن‌ها تحت فرض صفر غیر ممکن و یا دشوار است. در این مقاله، به بیان روشی برای آزمون چنین فرض‌هایی می‌پردازیم.

۲ روش مونت‌کارلو

شبیه‌سازی به دو صورت ایستا و پویا انجام می‌شود. در شبیه‌سازی ایستا زمان نقش ندارد در حالی که در شبیه‌سازی پویا، شبیه‌سازی در طول زمان انجام می‌شود. **روش مونت کارلو** از نظر شبیه‌سازی یک روش ایستا به‌حساب می‌آید و برای حل مسائلی از دو نوع «غیر تصادفی» مانند محاسبه‌ی انتگرال معین و «تصادفی» مانند استنباط آماری استفاده می‌شود؛ مرجع [۲] را ملاحظه نمایید. یکی از کاربردهای روش مونت کارلو، حل مسئله‌ی تصادفی آزمون فرض است که در این مقاله به آن پرداخته می‌شود؛ در مرجع [۱] و [۳]، کاربرد این روش در آزمون تصادفی کامل پیشامدها درالگوهای نقطه‌ای فضایی ارائه شده است.

[1] گروه آمار، دانشگاه یزد

دراین بخش، به بیان و بررسی روش مونت‌کارلو درآزمون فرض‌ها می‌پردازیم[۵]:

تحت فرض H_0، $m-1$ مشاهده‌ی مستقل از آماره‌ی آزمون U شبیه‌سازی می‌کنیم، $m=2,3,...$ ، و آن‌ها را $u_2,...,u_m$ می‌نامیم. مشاهده‌های مرتب شده را به صورت $u_{(1)},...,u_{(m)}$ نشان می‌دهیم؛ یعنی $u_{(1)} \leq ... \leq u_{(m)}$.

پیش از پرداختن به دنباله‌ی بحث، به بیان قضیه‌ای درباره‌ی آماره‌های ترتیبی می‌پردازیم[۱]:

قضیه ۱.۲ فرض کنید $U_1,...,U_m$ نمونه‌ای تصادفی از جامعه‌ای پیوسته با تابع چگالی احتمال f و تابع توزیع F باشد. اگر آماره‌های ترتیبی این نمونه را به صورت $U_{(1)},...,U_{(m)}$ نمایش دهیم، خواهیم داشت:

$$P(U_{(j)} \leq U_1) = \frac{m-j+1}{m}, \quad j=1,...,m$$

اثبات: با توجه به مفهوم احتمال شرطی داریم:

$$\begin{aligned}
P(U_{(j)} \leq U_1) &= \int_{-\infty}^{+\infty} P(U_{(j)} \leq U_1 | U_1 = u) f(u) du \\
&= \int_{-\infty}^{+\infty} P(U_{(j)} \leq u | U_1 = u) f(u) du \\
&= \int_{-\infty}^{+\infty} P(X \geq j-1) f(u) du,
\end{aligned}$$

که در آن متغیر تصادفی X نمایانگر تعداد U_iهایی است که کوچکتر یا برابر u هستند، $i=2,...,m$. با توجه به استقلال U_iها روشن است که $X \sim B(m-1, F(u))$.

بنابراین $P(U_{(j)} \leq U_1)$ برابر است با

$$\int_{-\infty}^{+\infty} \sum_{r=j-1}^{m-1} \binom{m-1}{r} [F(u)]^r [1-F(u)]^{m-1-r} f(u) du$$

$$\begin{aligned}
&= \sum_{r=j-1}^{m-1} \binom{m-1}{r} [(r!(m-1-r)!/m!] \\
&\times \int_{-\infty}^{+\infty} [m!/(r!(m-1-r)!][F(u)]^r \\
&\times f(u)[1-F(u)]^{m-1-r} du \\
&= \sum_{r=j-1}^{m-1} \binom{m-1}{r} [(r!(m-1-r)!/m!] \\
&= \sum_{r=j-1}^{m-1} \frac{1}{m} = \frac{m-j+1}{m}.
\end{aligned}$$

توجه شود که انتگرال رابطه‌ی دوم برابر یک است چون عبارت زیر انتگرال، تابع چگالی $(r+1)$امین آماره‌ی ترتیبی است.■

روشن است که در هر مسئله‌ی آزمون فرض، بسته به نوع فرض‌ها (یکطرفه یا دو طرفه بودن آن‌ها)، و همچنین آماره‌ی آزمون، ناحیه‌ی بحرانی فرق می‌کند. در این‌جا سه دسته‌ی مهم بررسی می‌شود:

دسته‌ی اول: اگر آماره‌ی آزمون به گونه‌ای باشد که با بزرگ بودن آن، انتظار رد فرض صفر را در برابر فرض مقابل داشته باشیم، منطقی است که اگر رتبه‌ی u_1، بزرگتر از k یی در مجموعه $N_m = \{1,...,m\}$ شود، باید فرض صفر را در برابر فرض مقابل رد کنیم و بنابراین ناحیه‌ی بحرانی آزمون، به صورت

$$C = \{u_1 | u_1 \geq u_{(k)}, \ni k \in N_m\},$$

می‌شود. اینک اگر بخواهیم اندازه‌ی آزمون برابر α باشد، باید داشته باشیم: $\alpha = \frac{m-k+1}{m}$. ولی در حالت کلی ممکن است برای هر α و m ای، k یی در N_m پیدا

نشود که در برابری بالا صدق کند؛ در حالی که اگر سطح آزمون را α فرض کنیم، k باید کوچکترین عدد صحیح متعلق به مجموعه‌ی N_m باشد که در نابرابری $k \geq m(1-\alpha)+1$ صدق کند. بنابراین

$$k = \lceil m(1-\alpha) \rceil + 1,$$

که در آن $\lceil x \rceil$، کوچکترین عدد صحیح ناکمتر از x یا سقف x است. در نتیجه، ناحیه‌ی بحرانی سطح α به صورت

$$C = \{u_1 | u_1 \geq u_{(\lceil m(1-\alpha) \rceil + 1)}\},$$

می‌شود.

دسته‌ی دوم : اگر فرض‌های صفر و مقابل و آماره‌ی آزمون به گونه‌ایی باشند که با بسیار کوچک شدن U، انتظار رد فرض صفر را در برابر فرض مقابل داشته باشیم، منطقی است که H_0 را در برابر H_1 رد کنیم هرگاه رتبه‌ی u_1 کوچکتر از k یی در N_m شود. در این حالت، ناحیه‌ی بحرانی سطح α، به صورت

$$C = \{u_1 | u_1 \geq u_{(k)}, \ni k \in N_m\},$$

می‌شود. با استدلالی شبیه حالت اول، می‌توان نشان داد که k باید برابر $\lfloor m\alpha \rfloor$ باشد. درنتیجه،

$$C = \{u_1 | u_1 \leq u_{(\lfloor m\alpha \rfloor)}\},$$

که در آن $\lfloor x \rfloor$ ، جزء صحیح x، یعنی بزرگترین عدد صحیح نابیشتر از x است.

دسته‌ی سوم: فرض‌های H_0 و H_1 و آماره‌ی آزمون را به گونه‌ایی در نظر بگیرید که با بسیار بزرگ یا کوچک شدن U، انتظار رد فرض صفر را در برابر فرض مقابل داشته باشیم. منطقی است که H_0 را در برابر H_1 رد کنیم هرگاه رتبه‌ی u_1 بسیار بزرگ یا بسیار کوچک شود. بنابراین،

$$\begin{aligned} C &= \{u_1 | u_1 \leq u_{(k_1)} \vee u_1 \geq u_{(k_2)}, \\ &\qquad \ni k_1, k_2 \in N_m, k_1 < k_2\}. \end{aligned}$$

در نتیجه اگر سطح آزمون را α فرض کنیم، باید داشته باشیم:

$$P(U_1 \leq U_{(K_1)}) + P(U_1 \geq U_{(K_2)}) \leq \alpha.$$

بنابراین، بنابر قضیه ۱.۲ و نابرابری بالا، k_1 و k_2 باید به ترتیب بزرگترین و کوچکترین عدد صحیح متعلق به N_m باشند که در نابرابری

$$\frac{k_1}{m} + \frac{m - k_2 + 1}{m} \leq \alpha,$$

و یا در هم‌ارز آن، یعنی

$$k_1 - k_2 \geq m(1-\alpha) - 1,$$

صدق کنند. به عنوان مثال، اگر $m = 100$ و $\alpha = 0.05$ ، باید داشته باشیم:

$$k_1 - k_2 \leq -96,$$

و بنابراین باید (k_1, k_2) برابر یکی از زوج مرتب‌های

$$(4,100), (3,100), (2,100), (1,100), (3,99), \\ (2,99), (1,99), (2,98), (1,98), (1,97),$$

باشد. ولی برای گزینش درست (k_1, k_2)، باید آگاهی کاملی درباره‌ی توزیع آماره‌ی U داشته باشیم. در حالت

ویژه، اگر دنباله‌ها را برابر بپنداریم، می‌توان نابرابری بالا را به صورت دو نابرابری $\frac{k_1}{m} \le \frac{\alpha}{2}$ و $\frac{\alpha}{2} \le \frac{m-k_2+1}{m}$ نوشت. بنابراین ناحیه‌ی بحرانی سطح α در این حالت به صورت

$$C = \{u_1 | u_1 \le u_{(\lfloor m\frac{\alpha}{2}\rfloor)} \vee u_1 \ge u_{(\lceil m(1-\frac{\alpha}{2})\rceil+1)}\},$$

در می‌آید. در نتیجه برای مثال بالا داریم: $k_1 = 2$ و $k_2 = 99$.

ممکن است در هر یک از حالت‌های گفته شده، برای برخی مقادیر m و α، ناحیه‌ی بحرانی تهی به‌دست بیاید. مثلاً با $m = 10$ و $\alpha = 0/05$، برای آزمون‌های حالت اول، ناحیه‌ی بحرانی تهی می‌شود؛ چون $\lceil m(1-\alpha)\rceil + 1$ برابر ۱۱ می‌شود که در N_{10} وجود ندارد.

برای انجام آزمون فرض‌های فوق می‌توان از «$p-$مقدار» نیز استفاده کرد. فرض کنید k برابر مقدار مشاهده شده رتبه u_1 در نمونه باشد. در آزمون فرض‌های دسته‌ی اول و دوم، $p-$ مقدار به ترتیب برابر $P(U_1 \ge U_{(k)}|H_0) = \frac{m-k+1}{m}$ و $P(U_1 \le U_{(k)}|H_1) = \frac{k}{m}$ به‌دست می‌آید. حال اگر $p-$ مقدار از α کوچکتر شود، H_0 را در برابر H_1 رد می‌کنیم. دقت شود که در دسته‌ی سوم $p-$مقدار تعریف نمی‌شود اما می‌توان معیار $p^* = 2\min\left(P(U_1 \ge U_{(k)}|H_0), P(U_1 \le U_{(k)}|H_0)\right)$ که برای برخی موارد از یک تجاوز می‌کند (به عنوان مثال در حالت $m = 3$، $k = 2$، مقدار p^* برابر $\frac{4}{3}$ به دست می‌آید.) را تعریف کرد. می‌توان نشان داد که همچون $p-$مقدار، اگر p^* کوچکتر از α شد، باید H_0 را رد کرد.

در دنباله به ذکر چند نکته‌ی مهم پیرامون روش مونت‌کارلو می‌پردازیم:

۱) در روش مونت‌کارلو بر خلاف روش‌های کلاسیک آزمون فرض، نیازی به دانستن توزیع آماره‌ی آزمون تحت فرض صفر نداریم تنها باید بتوان این آماره را تحت این فرض شبیه‌سازی کنیم.

۲) تعداد شبیه‌سازی‌ها $m - 1$، باید در حد امکان زیاد باشد؛ میزان بزرگی آن به ساختار مسئله آزمون فرض بستگی دارد. برای جزییات بیشتر [٦] را ملاحظه کنید.

۳ دو مثال کاربردی

مثال ۱.۳ فرض کنید $X_1, ..., X_n \sim N(\mu, 4)$. می‌خواهیم فرض ساده‌ی $H_0 : \mu = 0$ را در برابر فرض ساده‌ی $H_1 : \mu = 2$ آزمون می‌کنیم. می‌دانیم که بنابر لم نیمن—پیرسون، ناحیه‌ی رد آزمون MP، به صورت

$$C = \{\mathbf{x} = (x_1, \ldots, x_n) | \overline{x} > a\},$$

است؛ یعنی H_0 در برابر H_1 هنگامی رد می‌شود که مقدار $\overline{x}$ بسیار بزرگ باشد. در این مسئله، تحت H_0 داریم:

$$\overline{X} \sim N(0, \frac{4}{n}).$$

یعنی توزیع $\overline{X}$ تحت H_0 معلوم است. بنابراین می‌توان به آسانی آزمون خواسته شده را انجام داد و ناحیه‌ی بحرانی اندازه‌ی α را به صورت

$$C = \{\mathbf{x} = (x_1, ..., x_n) | \overline{x} > \frac{2}{\sqrt{n}} z_\alpha\},$$

به دست آورد؛ [۷] را ملاحظه کنید. اینک فرض کنید که بی‌توجه به آزمون فوق، ناشیانه میانگین مرتبه ۴، یعنی

$(\frac{1}{n}\sum {x_i}^4)^{\frac{1}{4}}$ را به جای میانگین حسابی، یعنی $\overline{x}$ به کار برده‌ایم. روشن است که توزیع آماره $U = (\frac{1}{n}\sum {x_i}^4)^{\frac{1}{4}}$ را نمی‌توان تحت H_0 به دست آورد. این جاست که به کارگیری روش مونت کارلو موجه جلوه می‌کند. فرض کنید مشاهده‌ی یک نمونه تصادفی به اندازه‌ی ۱۲ از توزیع بالا به صورت

$$\mathbf{x} = (-۱/۱۲, ۰/۹۵, ۰/۶۹, -۰/۰۹, -۰/۱۶, ۲/۲۶, ۱/۴۸, -۲/۲۹, ۱/۰۶, ۲/۰۳, ۱/۲۹, -۰/۹۳),$$

باشد. با توجه به این نمونه، میانگین حسابی و میانگین مرتبه ۴، به ترتیب برابر ۰/۴۳ و۱/۶۲ می‌شود. با توجه به ناحیه‌ی بحرانی MP، چون

$$0/43 < \frac{2}{\sqrt{12}} z_{0/05} = 0/95,$$

با اندازه‌ی آزمون ۰/۰۵، دلیلی برای رد H_0 در برابر H_1 وجود ندارد.

به منظور استفاده از روش مونت‌کارلو، با نرم افزار MINITAB، مستقلاً ۲۹ نمونه‌ی تصادفی هر یک به اندازه‌ی ۱۲ از توزیع $N(0, 4)$ شبیه‌سازی کرده‌ایم و ۲۹ مشاهده‌ی میانگین مرتبه‌ی ۴، به صورت زیر به دست آمده است:

$$(u_2, u_3, ..., u_{30}) = (۲/۰۱, ۲/۱۱, ۲/۴۰, ۲/۸۰, ۱/۶۵, ۲/۳۲, ۲/۴۳, ۲/۲۸, ۳/۱۲, ۲/۸۴, ۲/۶۴, ۴/۰۰, ۲/۳۵, ۳/۳۵, ۲/۰۶, ۲/۷۷, ۲/۸۹, ۱/۳۵, ۲/۷۰, ۱/۹۴, ۲/۶۴, ۱/۹۷, ۲/۷۴, ۲/۶۵, ۳/۶۵, ۲/۲۵, ۲/۷۷, ۱/۹۶, ۲/۰۵).$$

باافزودن $u_1 = ۱/۶۲$ به نمونه‌ی u_2،...، u_{30} و مرتب کردن آن‌ها، خواهیم داشت:

$$u_1 = u_{(2)},$$

ولی آماره‌ی U به گونه‌ای است که با بزرگ شدن مقدار آن، انتظار رد H_0 در برابر H_1 را داریم. بنابراین ناحیه‌ی بحرانی سطح ۰/۰۵ به صورت

$$C = \{u_1 | u_1 \geq u_{(30)}\} = \{u_1 | u_1 = u_{(30)}\},$$

می‌شود. درنتیجه چون ۱/۶۲ ≠ ۴/۰۰، در سطح ۰/۰۵ دلیلی برای رد H_0 در برابر H_1 نداریم.

$p-$ مقدار این آزمون برابر $0/97 = \frac{30-2+1}{30}$ به دست می‌آید؛ پس در سطح ۰/۰۵، دلیلی برای رد H_0 در برابر H_1 نداریم. دیده می‌شود که نتیجه دو آزمون کلاسیک و مونت کارلو یکی است.

اکنون اگر در همین مسئله بخواهیم، فرض ساده‌ی $H_0 : \mu = 0$ را در برابر $H_1 : \mu \neq 0$ بیازماییم، روشن است که میانگین مرتبه ۴، دوری از میانگین را چه در جهت مثبت و چه در جهت منفی، به یک اندازه می‌سنجد و بنابراین در این حالت نیز دلیلی برای رد H_0 در برابر H_1 وجود ندارد.

مثال ۲.۳ فرض کنید X_1، ...، X_n، نمونه‌ای تصادفی از جامعه‌ای با تابع چگالی احتمال

$$f(x;\alpha,\beta) = \frac{1}{2\beta}\sin(\frac{x-\alpha}{\beta}),\ \alpha < x < \alpha+\beta\pi, \quad \alpha \in \mathbb{R}, \beta \in \mathbb{R}^+$$

باشد. این توزیع که به خانواده‌ی توزیع‌های مکانی− مقیاسی تعلق دارد دارای میانگین $\alpha + \beta\frac{\pi}{2}$ و واریانس $\beta^2(\pi^2/4 - 2)$ است.

فرض کنید $\alpha = 0$. می‌خواهیم با داشتن مشاهده‌ی یک

نمونه‌ی ده تایی از این جامعه به صورت

$$(x_۱, ..., x_{۱۰}) = (۱/۷۳, ۰/۸۱, ۳/۰۵, ۱/۷۱, ۳/۷۶, ۴/۷۱, ۰/۶۳, ۱/۷۲, ۲/۴۶, ۱/۲۳),$$

فرض $H_۰ : \beta = ۱$ را در برابر فرض $H_۱ : \beta \neq ۱$ در سطح ۰/۰۵ آزمون کنیم.

با توجه به واریانس توزیع، منطقی است $H_۰$ را در برابر $H_۱$ رد کنیم هرگاه انحراف معیار نمونه‌ای، یعنی $\sqrt{\frac{\sum(X_i-\overline{X})^۲}{n-۱}}$، بسیار بزرگ یا بسیار کوچک شود در نتیجه برای این آزمون می‌توانیم آماره‌ی $\sqrt{\frac{\sum(X_i-\overline{X})^۲}{n-۱}}$ را به‌کار بریم. به‌دست آوردن توزیع این آماره تحت فرض صفر در حالت کلی غیر ممکن است. پس می‌توان برای این آزمون، روش مونت‌کارلو را به‌کار برد. مقدار $u_۱$ از نمونه فوق برابر ۱/۳۱ به‌دست می‌آید. با ۴۱ بار شبیه‌سازی U تحت فرض صفر، مشاهده‌های $u_۲$،...، $u_{۴۲}$ به صورت زیر به دست آمده است:

$$(u_۲, ..., u_{۴۲}) = (۰/۵۸, ۰/۵۴, ۰/۷۹, ۰/۵۶, ۰/۶۶, ۰/۷۳, ۰/۴۹, ۰/۵۹, ۰/۹۹, ۰/۶۹, ۱/۰۳, ۰/۶۳, ۰/۷۵, ۰/۶۱, ۰/۹۸, ۰/۷۱, ۰/۶۳, ۰/۶۰, ۰/۹۳, ۰/۴۵, ۰/۵۳, ۱/۰۶, ۰/۸۱, ۰/۷۳, ۰/۶۱, ۱/۰۳, ۰/۷۳, ۰/۶۷, ۰/۵۳, ۰/۶۸, ۰/۷۳, ۰/۸۱, ۰/۵۱, ۰/۶۱, ۰/۹۱, ۰/۷۶, ۱/۲۱, ۰/۷۰, ۰/۶۹, ۰/۶۸, ۰/۷۱).$$

به‌سادگی دیده می‌شود که رتبه‌ی $u_۱$ در بین داده‌های $u_۱$،...، $u_{۴۲}$ برابر ۴۲ است؛ یعنی $u_۱$ از همه‌ی مقادیر $u_۲, ..., u_{۴۲}$ بزرگتر است. از طرفی $m\frac{\alpha}{۲}$ و $m(۱-\frac{\alpha}{۲})+۱$ به ترتیب برابر ۱/۰۵ و ۴۱/۹۵ می‌شوند. بنابراین،

$$\begin{aligned} C &= \{u_۱ | u_۱ \leq u_{(۱)} \vee u_۱ \geq u_{(۴۲)}\} \\ &= \{u_۱ | u_۱ = u_{(۱)} \vee u_۱ = u_{(۴۲)}\}. \end{aligned}$$

حال چون $u_{(۴۲)} = ۱/۳۱$، درنتیجه با $\alpha = ۰/۰۵$ فرض $H_۰$ در برابر $H_۱$ رد می‌شود. برای این آزمون p^* مقدار برابر $۰/۰۴۸ = \frac{۴۲-۴۲+۱}{۴۲} \times ۲$ به دست می‌آید که کمتر از ۰/۰۵ است.

مراجع

[۱] ترابی، ح. (۱۳۷۵)، چند روش پیشنهادی برای آزمون CSR در الگوهای نقطه‌ای فضایی، پایان‌نامه کارشناسی ارشد.

[2] Banks, J. and Carson, J. S. (1984), *Discrete-Event System Simulation*, Prentice-Hall INC.

[3] Besage, J. and Diggle, P. J. (1977), *Simple Monte-Carlo Tests for Spatial Patterns*, Appl. Stat., 26, 327-333.

[4] Casella, G. and Berger, R. L. (2002), *Statistical Inference*, 2nd Edition, Duxbury Press.

[5] Hall, P. and Tilterington, D. M. (1989), *The Effect of Simulation Order on Level Accuracy and Power? Monte-Carlo Tests*, J. R. Stat. Soc. B, 51, 459-467.

[6] Hope, A. C. A. (1968), *A Simplified Monte-Carlo Significance Test Procedure*. J. R. Stat. Soc. B, 30, 582-598.

[7] Marriott, F. H. C. (1979), *Monte-Carlo Tests: How Many Simulations?* App. Stat. 28, 75-77.

[8] Robert, C. P. and Casella, G. (1999), *Monte Carlo Statistical Methods*, 2nd Ed., Springer, Berlin.

[9] Shao, J. (2003), *Mathematical Statistics*, Second Edition, Springer-Verlag, New York.