

دونکته از این معنی:

قضیه بهادر

سید محمود طاهری^۱

چکیده

در این مقاله، نخست مفهوم قضیه بهادر تشریح می‌شود. سپس روش‌های متفاوت اثبات آن ارایه و مقایسه می‌شوند. با چند مثال، توضیح می‌دهیم که عکس قضیه همواره برقرار نیست.

مقدمه

فرض بر این است که خوانتنده با مفاهیمی مانند آماره، آماره فرعی، بسنده، بسنده‌گی، بسنده‌گی مینیمال و کامل بودن آشنا می‌باشد.

۱. نکته اول:

مفهوم قضیه بهادر

نخست توضیحاتی درباره آماره بسنده، آماره بسنده مینیمال و آماره فرعی می‌دهیم. می‌دانیم که ویژگی یک آماره بسنده، این است که اطلاعاتی را که داده‌ها درباره پارامتر مجهول بیان می‌دارند، فشرده و درخود مرکز می‌کند. آماره بسنده‌ای که بیشترین فشردگی و کاهش در داده‌ها را انجام دهد، بسنده مینیمال نام دارد. به بیان ریاضی، آماره بسنده‌ای که تابعی از هر آماره بسنده دیگر باشد، آماره بسنده مینیمال است.

حال سؤال این است که: از دید شهودی چه ویژگی باعث

از زیباییهای ریاضیات، روش‌های متفاوت اثبات یک قضیه است. هر روش، پنجره‌ای به معنی و مفهوم پوشیده در قضیه می‌گشاید و نشان می‌دهد که آن قضیه چگونه با مباحث دیگر پیوند می‌خورد. بررسی روش‌های گوناگون اثبات یک قضیه، از دیدگاه آموزشی نیز حائز اهمیت است. به این ترتیب، عملاً نشان داده می‌شود که چگونه می‌توان از ابزارهای مختلف ریاضی استفاده کرده و آنها را برای رسیدن به یک هدف هماهنگ نمود.

در این مقاله به یکی از فضایی‌ای آمار ریاضی که نخستین بار توسط بهادر، آماردان هندی، ارایه و اثبات شد می‌پردازیم. در بخش اول معنا و مفهوم پنهان در قضیه تشریح می‌شود. در بخش دوم، سه اثبات برای آن بیان می‌کنیم. در بخش پایانی با چند مثال، نشان می‌دهیم که عکس قضیه همواره برقرار نیست.

^۱ سید محمود طاهری، دانشکده علوم ریاضی، دانشگاه صنعتی اصفهان (Taheri@cc.iut.ac.ir)

۲. نکته دوم:
اثبات‌های گوناگون برای قضیه بهادر
در ادامه فرض می‌کنیم که X_1, \dots, X_n ، نمونه‌ای تصادفی از جامعه‌ای با تابع چگالی احتمال $f(x; \theta)$ باشد.

روش اول: ۱) فرض کنید T آماره بسنده کامل و S آماره بسنده مینیمال باشد. ثابت می‌کنیم که T و S معادل هم و در واقع یک آماره‌اند. بدین معنی که T تابعی از S است، و S تابعی از T .

بدیهی است که S تابعی از T باشد، زیرا آماره بسنده مینیمال تابعی از هر آماره بسنده است. کافی است ثابت کنیم T تابعی از S است.

تابع $f(T)$ را به صورت زیر در نظر بگیرید:

$$f(T) = T - E(T|S) = T - g(S)$$

اما $E[f(T)] = 0$ ، پس بنایه کامل بودن T داریم: $P_\theta[T = g(S)] = 1 \forall \theta$ و یعنی $P_\theta[f(T) = 0] = 1 \forall \theta$

روش دوم: فرض کنید T آماره بسنده کامل باشد، ثابت می‌کنیم T تابعی از هر آماره بسنده دلخواه S است ولذا مینیمال است. دو تابع $h(S)$ و $L(T)$ را به صورت زیر تعریف می‌کنیم:

$$h(S) = E(T|S)$$

$$L(T) = E[h(S)|T]$$

داریم

$$E[L(T)] = E[h(S)] = E[T]$$

و بنایه قضیه رائو-بلکول

$$Var[L(T)] \leq Var[h(S)] \leq Var(T) \quad (1)$$

از طرف دیگر $E[L(T) - T] = 0$ و چون T کامل است پس $Var[L(T)] = Var(T)$ و لذا $L(T) = T$

می‌شود و تضمین می‌کند که یک آماره، بیشترین فشردگی در داده‌ها را متحمل (متقبل؟!) شود؟ جواب به این سؤال پیوند نزدیک با مفهوم آماره فرعی (کمکی) دارد. ابتدا اشاره می‌کنیم که یک آماره، در صورتی می‌تواند به ما درباره پارامتر مجھول آگاهی ببخشد که توزیع آماره، به آن پارامتر بستگی داشته باشد. اگر آماره‌ای اینگونه نباشد، آماره فرعی نامیده می‌شود [۶]. به همین جهت آماره‌ای را که فرعی نباشد، آگاهی بخش می‌نماید. بدیهی است که اگر T یک آماره فرعی و (T) تابعی از آن باشد، آنگاه $E_\theta[g(T)]$ مقداری ثابت است و به θ بستگی ندارد. آماره‌ای که ویژگی اخیر را (برای هر تابع دلخواه $(g(t))$ داشته باشد، آماره فرعی مرتبه اول نامیده می‌شود).

اگر نون فرض کنید T یک آماره بسنده باشد. اگر بتوان تابعی از T ساخت که آماره فرعی باشد، این بدین معنی است که: «در T آماره فرعی وجود دارد» و به عبارت دیگر: «هنوز در T چیزهایی هست که اطلاعی درباره θ به ما نمی‌دهد» ولذا می‌توان نتیجه گرفت که: « T هنوز کاملاً فشرده نشده است». با این توضیحات، آشکار می‌شود که یک آماره بسنده T در بیان و فشردگی داده‌ها کاملاً تواناست، اگر هیچ تابعی (غیر از تابع ثابت) از T یافت نشود که آماره فرعی، و در واقع حتی فرعی مرتبه اول، باشد. یعنی اگر $E_\theta[g(T)] = 0 \forall \theta$ ، که با کم کردن ثابت C از طرفین به شرط زیر می‌رسیم:

$$E_\theta[g(T)] = 0 \forall \theta \Rightarrow P_\theta[g(t) = 0] = 1 \forall \theta$$

و این چیزی نیست جز تعریف کامل بودن یک آماره [۱۰]. پس کامل بودن یک آماره بسنده، به طور ذاتی و اساسی منجر به بیشترین فشردگی ممکن در داده‌ها می‌شود و لذا تعجب ندارد اگر ادعا شود که:

قضیه بهادر: هر آماره بسنده کامل،
بسنده مینیمال است.

(دقیق) کنید که پیشامد $c_i = T(X)$ زیرمجموعه پیشامد $T^*(X) = a$ می‌باشد). پس برای هر θ ، $E_\theta h(T) = 0$. اما طبق فرض برای دست‌کم یک θ ، $P_\theta[T^*(X) = a] > 0$ و لذا مثلاً $P_\theta[T(X) = c_1] > 0$ ، که در این حالت $h(T(X)) = h(c_1)$ و در نتیجه T کامل نیست.

مقایسه سه روش اثبات: روش اول، ساده‌تر از سایر روش‌هاست و از این نظر بهترین است. در روش دوم، از شیوه زیبایی استفاده شده است که همانا بکارگیری موقت (کاتالیزور وار) تابع $L(T)$ است؛ به منظور سود جستن از کامل بودن T و استفاده از قضیه رائو-بلکول. البته طولانی بودن و اتکاء به قضیه رائو-بلکول می‌تواند دو نقطه ضعف این روش تلقی شود. روش سوم کمی پیچیده است و از نظر سادگی و روانی به روش دوم و به ویژه به روش اول نمی‌رسد. اما دو نکته است که مزیت روش سوم می‌باشد:

الف) در روش سوم فرض نمی‌شود که آماره بسنده مینیمال وجود دارد. اما در روش اول باید وجود آنرا فرض کنیم. گرچه تحت شرایطی ضعیف آماره بسنده مینیمال همیشه وجود دارد [۱۲-۴]، ولی به هر حال این یک پیش‌فرض است که در روش سوم احتیاجی به آن نیست.

ب) روش سوم تقریباً یک روش ساختاری است. به این بیان که در آن، مستقیماً از تعریف کامل بودن استفاده شده و تابعی که در شرط دلخواه صدق می‌کند ساخته می‌شود. این، به ویژه از دید ریاضیدان شهودگرا، یک مزیت است. (البته از دید چنین ریاضیدانی پای استدلال سوم چویین است: شهودگرایان، و به تعبیر دیگر ساختارگرایان، به معادل بودن گزاره $A \sim B \Rightarrow A \sim B$ اعتقاد ندارند!)

۳. چند بحث تکمیلی

۱.۳ مینیمال بودن، به خودی خود تضمین نمی‌کند که یک آماره شامل اطلاعات فرعی و غیر لازم درباره θ نباشد (کامل

پس نابرابریها در (۱) در واقع برابری هستند یعنی T و $h(S)$ و $L(T)$ سه براوردگر ناریب ($E_\theta(T)$ می‌باشند که واریانس‌های یکسان دارند؛ به ویژه $Var[h(S)] = Var(T)$). اما مجدداً بنابر قضیه رائو-بلکول، برابری واریانس‌های دو براوردگر T و $h(S)$ هنگامی است که این دو براوردگر با احتمال یک برابر باشند، یعنی $(S) = h(S)T$. یعنی آماره بسنده کامل T تابعی از آماره بسنده دلخواه S است.

یادآور می‌شویم که اثبات نابرابریها (۱) و این موضوع را که برابری واریانس‌های T و $h(S)$ نتیجه می‌دهد که $T = h(S)$ می‌توان مستقیماً برایه نامساوی گشتاوری یعنی آنچه اساس قضیه رائو-بلکول است، بدست آورد [۲].

روش سوم: [۷] ثابت می‌کنیم هر آماره بسنده که مینیمال نباشد، کامل نخواهد بود.

برای این، فرض کنید T یک آماره بسنده باشد ولی مینیمال نباشد. تابعی مانند $g(T) \neq 0$ می‌سازیم که امید ریاضی آن صفر است.

ساختن تابع $h(T)$ با شرط فوق: چون آماره T مینیمال نیست پس آماره بسنده T^* و تابعی مانند g وجود دارد که $T^*(X) = g[T(X)]$ به قسمی که برای دست‌کم یک a و دو مقدار c_1 و c_2 داشته باشیم:

$$g(c_1) = g(c_2) = a$$

(برای دست‌کم یک θ)

$$P[T(X) = c_i | T^*(X) = a] > 0 \quad i = 1, 2$$

اکنون تابع h را اینگونه تعریف می‌کنیم:

$$h(c) = \begin{cases} \frac{(-1)^i}{P_\theta[T(X) = c_i | T^*(X) = a]} & c = c_i, i = 1, 2 \\ 0 & \text{سایر نقاط} \end{cases}$$

داریم:

$$\begin{aligned} E[h(T(X))] &= \sum_{i=1}^2 \frac{(-1)^i}{P_\theta[T(X) = c_i | T^*(X) = a]} \\ &\times P[T(X) = c_i] \\ &= (-1 + 1) \times P[T^*(X) = a] \\ &= 0 \end{aligned}$$

مثال ۲. فرض کنید X_1, \dots, X_n نمونه‌ای تصادفی از توزیع یکنواخت $(\frac{1}{n}, \theta + \frac{1}{n})$ باشد. در اینجا $T(X) = (X_{(1)}, X_{(n)})$ آماره بستنده مینیمال است اما کامل نیست، زیرا مثلاً $E_\theta[X_{(n)} - X_{(1)}] = \frac{n-1}{n+1}$. حتی در حالت $n=1$ نیز X کامل نیست. جالب آنکه در این حالت X ، برآورده‌گر ناریب با کمترین واریانس (UMVUE) برای θ نیست و اصولاً در این حالت UMVUE برای θ وجود ندارد. برای جزئیات به [۷] صفحه ۱۶۶ مراجعه کنید.

۲۰۳ بحثی که در بخش اول ارایه کردیم به فهم شهودی قضیه باسو نیز کمک می‌کند. چنانچه آماره‌ای بستنده کامل باشد، بدین معنی است که علاوه بر داشتن همه اطلاعات لازم درباره θ ، هیچ اطلاعی بیشتر از آنچه درباره θ لازم است ندارد. پس چنین آماره‌ای ارتباطی با یک آماره فرعی (که شامل هیچ اطلاعی درباره θ نیست) نمی‌تواند داشته باشد. از این رو طبیعی به نظر می‌رسد که هر آماره فرعی مستقل از یک آماره بستنده کامل باشد.

درباره آماره کامل و قضیه باسو و مباحث مریبوط به آنها، مقاله [۹] حاوی نکات ارزنده‌ای است. همچنین در مقاله جالب و روان [۳] مباحث مختلف مریبوط به آماره فرعی بررسی شده است که برای خواننده علاقمند، می‌تواند مفید باشد.

بودن چنین است). مثالهای فراوانی وجود دارند که می‌توانیم از یک آماره بستنده مینیمال، تابعی بسازیم که آماره فرعی باشد، یعنی توزیع آن به θ بستگی نداشته باشد. این مثالها نشان می‌دهند که:

لزوماً هر آماره بستنده مینیمال، کامل نیست ولذا عکس قضیه بهادر همواره برقرار نیست. نتیجه آنکه، بستنده کامل بودن یک شرط قوی‌تر از بستنده مینیمال بودن است و البته در استنباط آماری یک مفهوم بسیار کارساز می‌باشد [۱۰ و ۱۱]. برای توضیح مطلب، در ادامه چند مثال ارایه شده است که در آنها آماره بستنده مینیمال، کامل نیست.

مثال ۱. سکه نالریبی را θ بار می‌ریزیم و $\{\theta, 1, 2, \dots\}$ آسانی می‌توان تحقیق کرد که X تعداد شیرها، یک آماره بستنده مینیمال برای θ است. اما تابع زیر که امید ریاضی آن همواره صفر است، نشان می‌دهد که X کامل نیست.

$$T(X) = \begin{cases} 1 & \text{اگر } X \text{ زوج باشد} \\ 0 & \text{اگر } X \text{ فرد باشد} \end{cases}$$

مثال ۲. فرض کنید X_1, \dots, X_n نمونه‌ای تصادفی از توزیع لجستیک زیر باشد

$$f(x; \theta) = \frac{e^{-(x-\theta)}}{[1 + e^{-(x-\theta)}]^2}, \quad \theta \in R$$

در اینجا (X_1, \dots, X_n) آماره بستنده مینیمال است، در حالیکه مثلاً $S(T) = X_{(n)} - X_{(1)}$ یک آماره فرعی است.

مراجع

- [۱] بهمودیان، جواد؛ آمار ریاضی، انتشارات امیرکبیر، ۱۳۷۰
- [۲] پارسیان، احمد؛ مبانی آمار ریاضی، انتشارات دانشگاه صنعتی اصفهان، ۱۳۷۸
- [۳] سیاره، عبدالرضا؛ آماره‌های کمکی و نقش آنها در استنباط آماری، اندیشه آماری، سال سوم، شماره اول، ۱۳۷۷
- [4] Bahadur, R. R. (1954) *Sufficiency and Statistical Decision Functions*, Ann. Math. Stat. 25: 423-462.
- [5] Bahadur, R. R. (1957) *On Unbiased Estimates of Uniformly Minimum Variance*, Sankhya 18: 211-224.

- [6] Basu, D. (1955) *On Statistics Independent of a Complete Sufficient Statistic*, Sankhya, 5: 377-380.
- [7] Kiefer, J. C. (1987) *Introduction to Statistical Inference*, Springer.
- [8] Lehman, E. L. and Scheffe, H. (1950) *Completeness, Similar regions, and Unbiased Estimation*, Sankhya 10: 305-340.
- [9] Lehman, E. L. (1981) *An Interpretation of Completeness and Basu's Theorem*, JASA, 76: 335-340.
- [10] Lehmann, E. L. (1983) *Theory of Point Estimation*, John Wiley.
- [11] Lehmann, E. L. (1991) *Testing Statistical Hypotheses*, John Wiley.
- [12] Pitcher, T. S. (1957) *Sets of Measures not Admitting Necessary and Sufficient Statistics or Subfields*, Ann. Math. Stat. 26: 267-268.
- [13] Rohatgi, V. K. (1976) *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley.