اثبات پذیرفتنی بودن آزمون‌های T^2 هتلینگ و R^2 به شیوهٔ بیزی

سید یاسر صمدی[۱] ، محمد رضا مشکانی[۱]

چکیده:

در مکتب بسامدگرای استنباط آماری یکی از خواص مطلوب برای شیوه‌های آماری خاصیت پذیرفتنی بودن آنها است. اثبات پذیرفتنی بودن یک شیوهٔ آماری به جز در حالت‌های ساده، کار چندان آسانی نیست. از سوی دیگر، شیوه‌های بیزی اغلب پذیرفتنی‌اند. مخصوصاً اگر مخاطره بیزی یک شیوه بیزی نسبت به توزیع پیشینی (سره یا ناسره) متناهی باشد، آن شیوه پذیرفتنی است. بنابراین، راهی دیگر برای اثبات پذیرفتنی بودن یک شیوهٔ استنباطی آن است که نشان داده شود، آن شیوه یک شیوهٔ بیزی متناظر با یک توزیع پیشینی بوده و دارای مخاطره متناهی است. در این مقاله، به منظور توضیح این روش اثبات پذیرفتنی بودن شیوه‌های استنباطی بسامدگرا، از شیوهٔ بیزی به عنوان یک ابزار اثبات استفاده می‌کنیم تا پذیرفتنی بودن شیوهٔ آزمون فرض در توزیع نرمال چندمتغیره را با T^2 هتلینگ و R^2 (ضریب همبستگی چندگانه) نشان دهیم. ارزش روش بیزی وقتی آشکار خواهد شد که به اثبات پذیرفتنی بودن این شیوه‌ها به روش مستقیم بیندیشیم. یعنی نشان دهیم که مخاطرهٔ آنها به ازای همه مقادیر پارامتر از مخاطرهٔ سایر شیوه‌ها کمتر یا مساوی و دست‌کم به ازای برخی از مقادیر پارامتر اکیداً کمتر است.

واژه‌های کلیدی: آزمون فرض، پذیرفتنی بودن، شیوهٔ بیزی، شیوهٔ بسامدگرا، مخاطره، مخاطرهٔ بیزی.

۱ مقدمه

توانمندی روش‌های بیزی در حل مسائل آماری بر کسی پوشیده نیست. بخصوص در به‌دست‌آوردن رده برآوردگرهای بهینه، شیوهٔ بیزی، شیوه‌ای مطمئن، آسان و کاراتر نسبت به روش‌های دیگر و گاهی تنها روش ممکن برای دستیابی به برآوردگر بهینه است. در مواردی که حل مسئله به روش کلاسیک بسیار پیچیده و طولانی می‌شود با کمک شیوه بیزی می‌توان پیچیدگی و دشواری حل مسئله را تا حدود زیادی برطرف نمود و به راه‌حلی کوتاه دست یافت. مثلاً اگر بخواهیم به روش‌های کلاسیک نشان دهیم که آزمون‌های آماری کلاسیک پذیرفتنی‌اند، باید راه طولانی و دشواری را طی نماییم. ولی به شیوهٔ بیزی و با تعیین پیشین‌های مناسب می‌توان به راحتی این مهم را انجام داد. اکنون با معرفی آزمون‌هایی برای مسائل نرمال چند متغیره کلاسیک که اثبات پذیرفتنی بودن آنها به روش‌های کلاسیک کار سخت و پیچیده‌ای است، نشان داده می‌شود که پذیرفتنی بودن آنها با کمک شیوه بیزی نسبتاً آسان است. اساس کار در این نحوهٔ اثبات متکی به قضیه زیر است:

[۱] گروه آمار دانشگاه شهید بهشتی

قضیه ۱ اگر برآوردگر δ^π، متناظر با پیشین π (سره یا ناسره) یکتا باشد، δ^π پذیرفتنی است [۱۱].

توجه داریم که در روش بیزی آزمون‌های فرض نیز در قالب مسئلهٔ برآورد ریخته می‌شوند وفرضی که احتمال پسین بیشتری داشته باشد، فرض قابل قبول است. بنابراین لازم است احتمال پسین فرض‌های رقیب برآورد شوند، یعنی در هر حال با مسئلهٔ برآورد مواجه‌ایم. اساس بسیاری از تحلیل‌های آماری چند متغیرهٔ پیوسته، توزیع نرمال چند متغیره است. چگالی آن برای یک بردار p−بعدی $\mathbf{X}$ با میانگین μ و ماتریس کوواریانس $\mathbf{\Sigma}$ عبارت است از :

$$f(x \mid \theta) \propto |\Sigma|^{-\frac{1}{2}} \exp\{-\frac{1}{2}tr\Sigma^{-1}(x-\mu)(x-\mu)'\}$$

$$x \in R^p \quad \mu \in R^p \quad \Sigma \in \Omega \qquad (1)$$

که در آن $\theta \in (\mu, \mathbf{\Sigma})$ نشانگر مجموعه پارامترهای توزیع است و Ω مجموع ماتریس‌های مربع $p \times p$ همیشه مثبت است. پذیرفتنی بودن آزمون‌های آماری کلاسیک با استفاده از یکی از شیوه‌های زیر اثبات می‌شود:

(۱) شیوه‌های بیزی

(۲) استفاده از ساختار ویژه خانواده نمایی

(۳) ناوردایی

(۴) استفاده از ویژگی‌های موضعی.

مثال‌هایی از مورد (۱) را می‌توان در لی‌من و استاین[۸]، کارلین[۵]، لی‌من[۷] و الیسون[۳] پیدا کرد. شیوهٔ (۲) توسط برن‌بام[۲]، استاین[۱۳]، ناندی[۱۰]، گوش[۴] و شوارتز[۱۲] استفاده شده است. علاوه بر حالت بدیهی گروه‌های فشرده، تنها حالت انتقال پارامتر یک بعدی از مورد (۳) توسط لی‌من و استاین[۹] مطالعه شده است.

بیشترین استفاده از (۴) در حالت‌هایی است که نااریبی، تشابه را بر روی قسمتی از مرز $\mathbf{H}_0$ و $\mathbf{H}_1$ نتیجه می‌دهد. آنگاه تواناترین آزمون‌های نااریب، یکنواخت یا موضعی یکتا هستند (لی‌من[۷]). تواناترین آزمون‌های یکنواخت را می‌توان به عنوان حالت خاصی از (۱) تلقی کرد. برای مثال، می‌توان نتیجه‌ای همانند نتیجه‌ای را که والد[۱۴] راجع به آزمون تحلیل واریانس بدست آورد به عنوان کاربردی از (۱) برای آزمون‌های مشابه بدست آمده از نااریبی با توجه به (۴) در نظر گرفت.

استفاده از تکنیک‌های (۲)، (۳) و (۴) در مسائل نرمال چند متغیره استاندارد، محدود است. لازم نیست که بهترین شیوه‌های ناوردا تحت گروه خطی کامل مینیماکس باشد، بلکه فقط پذیرفتنی باشد، کافی است. برای اثبات وجود آزمون‌های پذیرفتنی، نمی‌توان از تواناترین آزمون‌های یکنواخت نااریب و یا به طور مشابه از قضیه والد استفاده کرد. در حالی که ساختار نمایی می‌تواند برای اثبات پذیرفتنی بودن آزمون $\mathbf{T}^2$ هتلینگ (استاین[۱۳]) و رده‌ای از آزمون فرض‌های خطی چند متغیره عام (شوارتز[۱۲]) استفاده گردد. از این خانواده (یا کلاً برای خانواده‌های غیرنمایی مشخص) نمی‌توان در مسئله آزمون استقلال مجموعه‌هایی از متغیرها وقتی که $p \geq 3$ است، استفاده کرد. این موضوع توسط استاین[۱۳] بحث شده است. از دیدگاهی اندکی متفاوت می‌توان دید که این روش حتی در حالت دو متغیره به شکست منجر می‌شود. چون $\frac{f_{R^2}(r^2, \rho^2)}{f_{R^2}(r^2, 0)}$ (که در آن R و ρ به ترتیب ضرایب همبستگی نمونه و جامعه هستند)، وقتی که $\rho \to 1$ کراندار نیست. پذیرفتنی بودن آزمون

معمولی در حالت دو متغیره یعنی $p = ۲$ ، در لی‌من [۷] با استفاده از (۴) اثبات شده است. اما این روش نیز برای $p > ۲$ به شکست منجر می‌شود.

حتی در حالتی مانند $\mathbf{T}^۲$ هتلینگ که پذیرفتنی بودن آن با روش (۲) اثبات شده است، شیوه بیزی اطلاعات بیشتری را برای انجام آزمون در اختیار می‌گذارد. با استفاده از روش (۲) تنها مطمئن هستیم که آزمون دیگری با همان اندازه وجود ندارد که بهتر از $\mathbf{T}^۲$ در جاهایی دور از $\mathbf{H}_\circ$ باشد. درحالی که نتیجهٔ بیزی رفتار $\mathbf{T}^۲$ را در نزدیکی $\mathbf{H}_\circ$ منعکس می‌کند.

با این حال تکنیک بیزی محدودیت‌هایی دارد. چون همیشه حدس پیشین Π که آزمون مفروض نسبت به آن، آزمونی بیزی باشد، یا برای یک پیشین مفروض با انتگرال قابل حلی مواجه شویم، ممکن است مشکل باشد. بعلاوه با این روش نمی‌توان اثبات کرد که بعضی از آزمون‌های پذیرفتنی طبیعی به خاطر دلایلی غیر از عدم انتگرال‌پذیری، برای مینیمم اندازه‌های نمونه، پذیرفتنی‌اند. به عنوان مثال تفسیر برن‌بام نشان می‌دهد که برای مسئله آزمون اینکه میانگین یک بردار نرمال دو متغیره با ماتریس کوواریانس معلوم صفر است، هر ناحیه ردی به شکل چندضلعی محدب فشرده پذیرفتنی است، اما به صورت تحلیلی نمی‌توان نشان داد که این یک آزمون بیزی است.

بنابراین همانطور که گفته شد اثبات پذیرفتنی بودن یک آزمون به شیوه‌های کلاسیک نه تنها کار آسانی نیست بلکه در برخی موارد امکان‌پذیر نیست. اکنون با ارایه دو آزمون معروف $\mathbf{T}^۲$ و $\mathbf{R}^۲$ و بیان خلاصه‌ای از اثبات پذیرفتنی بودن آنها به شیوه‌های کلاسیک، که بسیار سخت وپیچیده است نشان داده می‌شود که به شیوه بیزی و با تعیین پیشین‌های مناسب، خیلی آسان‌تر و با راه‌حل کوتاه‌تری پذیرفتنی بودن آنها ثابت می‌شود و توانایی این شیوه بیش از پیش آشکار می‌شود.

۲ نمادگذاری

در هر مسئله فضای پارامتر را با $\Theta = \{\theta\} = \mathbf{H}_\circ + \mathbf{H}_۱$ نشان خواهیم داد،(که درآن $\Theta = \mathbf{H}_\circ + \mathbf{H}_۱$ تأکید بر تقارن دو فرض دارد) و پارامتر θ شامل مجموعه‌ای از بردارهای μ و ماتریس‌های Σ و ... است. همواره Σ از مرتبه $p \times p$ یک ماتریس کوواریانس همیشه مثبت فرض می‌شود. اندازه‌های احتمال پیشین را با Π نشان می‌دهیم. نیازی نیست که $\Pi(\Theta) = ۱$ باشد، تنها کافی است $\Pi(\Theta) < \infty$ باشد. بنابراین برای راحتی در نوشتن از دادن مقادیر صریح مضرب‌های ثابت مثبت در Π خودداری می‌کنیم. اگر $\Pi = \Pi_\circ + \Pi_۱$ و Π_i یک اندازه متناهی برای $\mathbf{H}_i$، $i = ۱,۲$ باشد، آنگاه هر ناحیه بحرانی بیزی (متناظر با تابع زیان $۱ - \circ$) به فرم زیر است. به عبارت دیگر آزمون به شیوه بیزی (با تابع زیان $۱ - \circ$) فرض $\mathbf{H}_\circ$ را رد می‌کند، اگر:

$$\frac{\int f(x \mid \theta)\Pi_۱(d\theta)}{\int f(x \mid \theta)\Pi_\circ(d\theta)} \geq c\,, \qquad o \leq c \leq \infty \qquad (۲)$$

در همه مثال‌های ارائه شده برقراری برابری در رابطه (۲) به ازای هر $\theta \in \Theta$ با احتمال صفر رخ می‌دهد. بنابراین شیوه‌های بیزی ذاتاً یکتا هستند. در نتیجه همه شیوه‌های بیزی پذیرفتنی‌اند. توزیع‌های $\Pi_i(\theta)$، $i = ۱,۲$ که در مثال‌ها به کارمی‌روند، همگی دارای چگالی‌هایی بر

حسب اندازه لبگ روی مجموعه‌های اقلیدسی هستند. گاهی اوقات در نظرگرفتن Π_i یا یکی از عامل‌های آن که چگالی‌ای روی یک مجموعهٔ اقلیدسی Ψ داشته باشد، مناسب‌تراست. مجموعهٔ Ψ می‌تواند زیرمجموعه‌ای از Ω باشد که با نگاشتی معین تعریف می‌شود. به عنوان مثال محاسبات با چگالی لبگ $c|\mathbf{I}_p + \eta\eta'|^{-m}$ روی فضای اقلیدسی p، یعنی در فضای $\{\eta\} = \Psi = \mathbf{E}^p$ نسبت به محاسبات با اندازه تولید شده روی فضای ماتریس‌های همیشه مثبت $(\mathbf{I}_p + \eta\eta') = \Sigma$ آسان‌تر است. چنین چگالی‌هایی از نوع لبگ (که کافی است انتگرال پذیر باشند ونیازی نیست که انتگرال یک داشته باشند) با $\frac{d\Pi_i(\eta)}{d\eta}$ و انتگرال‌گیر اندازه لبگ با $d\eta$ نشان داده می‌شوند. در مثال‌های ارایه شده، همه چگالی‌ها روی فضای حاصل ضرب پارامتر و نمونه، پیوسته‌اند و بنابراین اندازه‌پذیرند.

روشن است که در بسیاری از مثال‌ها، به تعداد خیلی زیاد و اغلب نامتناهی Π مستقل خطی بسته به اینکه شیوه مفروض شیوه‌ای بیزی باشد، وجود دارد. برای مثال Π_i همه اندازه را به یک مجموعه که در آن $\mathbf{\Sigma} = \mathbf{B} + \eta\eta'$ و $\mathbf{B}$ یک ماتریس همیشه مثبت دلخواه (که معمولاً برابر با $\mathbf{I}_p$ می‌گیریم) و η یک ماتریس تصادفی $p \times q$ است، اختصاص می‌دهد. درمثال‌های ارائه شده میانگین‌ها و ماتریس‌های کوواریانس نامعلوم‌اند. در مواردی که بعضی از این پارامترها معلوم باشند، مسائل آسان‌تر حل می‌شوند و بعضی اوقات حل‌های پیش‌پاافتاده‌ای دارند.

۳ نتایج اولیه

اکنون برخی نتایج انتگرال گیری را که به طور مکرر استفاده خواهد شد، به طور خلاصه بیان می‌کنیم. اگر β برداری از مرتبه $p \times ۱$ و $\mathbf{C}$ یک ماتریس همیشه مثبت از مرتبه $p \times p$ باشد، آنگاه طبق رابطه (۱) داریم:

$$\int \exp\ tr\{-\frac{۱}{۲}(\Sigma\beta\beta' - ۲C\beta')\}\ d\beta = |\Sigma|^{-\frac{۱}{۲}}\ \exp\ tr\{\frac{۱}{۲}\Sigma^{-۱}CC'\} \qquad (۳)$$

اگر η یک بردارp−بعدی باشد، آنگاه می‌توان نشان داد که:

$$|I_p + \eta\eta'| = ۱ + \eta'\eta \qquad (۴)$$

چون داریم:

$$\begin{bmatrix} I_p + \eta\eta' & \eta \\ ۰ & ۱ \end{bmatrix} \begin{bmatrix} I_p & ۰ \\ -\eta' & ۱ \end{bmatrix} = \begin{bmatrix} I_p & ۰ \\ -\eta' & ۱ \end{bmatrix} \begin{bmatrix} I_p & \eta \\ ۰ & ۱ + \eta'\eta \end{bmatrix} \qquad (۵)$$

زیرا

$$\begin{bmatrix} I_p + \eta\eta' - \eta\eta' & \eta \\ -\eta' & ۱ \end{bmatrix} = \begin{bmatrix} I_p & \eta \\ -\eta' & ۱ + \eta'\eta - \eta'\eta \end{bmatrix}$$

بنابراین با گرفتن دترمینان از طرفین (۵)، رابطه (۴) بدست می‌آید. همچنین می‌توان نشان داد که:

$$\eta'(I + \eta\eta')^{-۱}\eta + (۱ + \eta'\eta)^{-۱} = ۱ \qquad (۶)$$

برای اثبات رابطه بالا می‌توان از رابطه زیر استفاده کرد:

$$(I + \eta\eta')^{-۱} = I - (۱ + \eta'\eta)^{-۱}\eta\eta' \qquad (۷)$$

اما درستی رابطه (۷) را می‌توان با ضرب طرفین در$(\mathbf{I} + \eta\eta')$ به صورت زیر نشان داد.

$$I = (I + \eta\eta') - \frac{(I + \eta\eta')\eta\eta'}{۱ + \eta\eta'}$$

اکنون η' را از راست و η را از چپ در رابطه جدید ضرب می‌کنیم. آنگاه:

$$\begin{aligned}\eta'\eta = (\eta'\eta)^2 - \frac{(\eta'\eta)^2-(\eta'\eta)^3}{1+\eta'\eta} &= \frac{\eta'\eta(1+\eta'\eta)}{1+\eta'\eta} \\ &= \eta'\eta\end{aligned}$$

سپس با جایگذاری رابطه (۷) در(۶) درستی آن محقق می‌شود. برای بردار$p-$بعدی η و هر عدد حقیقی k تعریف می‌کنیم:

$$J_k := \int |I+\eta\eta'|^{-\frac{k}{2}} d\eta$$

آنگاه

$$J_k<\infty \iff k>p \tag{8}$$

چون طبق رابطه (۴) داریم:

$$\begin{aligned} J_k &= \int (1+\eta\eta')^{-\frac{k}{2}} d\eta \\ &= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\left(\frac{1}{1+\sum{\eta_i}^2}\right)^{\frac{k}{2}} d\eta_1\dots d\eta_p\end{aligned}$$

تبدیل $\eta_j = \omega_j\sqrt{1+\sum_{i=j+1}^p {\eta_i}^2}$، $j=1,\dots,p-1$ را اعمال می‌کنیم آنگاه:

$$\begin{aligned}
\eta_{p-1} &= \omega_{p-1}\sqrt{1+\eta_p^2} \\
\eta_{p-2} &= \omega_{p-2}\sqrt{(1+\eta_p^2)(1+\omega_{p-1}^2)} \\
\eta_{p-3} &= \omega_{p-3}\sqrt{(1+\eta_p^2)(1+\omega_{p-1}^2)(1+\omega_{p-2}^2)} \\
&\vdots \\
\eta_1 &= \omega_1\sqrt{(1+\eta_p^2)\prod_{i=2}^{p-1}(1+\omega_i^2)}
\end{aligned}$$

ژاکوبی تبدیل عبارت است از:

$$\begin{aligned}\left|\frac{\partial(\eta_{p-1},\eta_{p-2},\dots,\eta_1)}{\partial(\omega_{p-1},\omega_{p-2},\dots,\omega_1)}\right| = (1+\eta_p^2)^{\frac{p-1}{2}} &\times \\ \prod_{i=2}^{p-1}(1+\omega_i^2)^{\frac{p-(p-i+1)}{2}}&\end{aligned}$$

از طرفی داریم:

$$\eta_1^2 = \omega_1^2\left(1+\sum_{i=2}^p\eta_i^2\right) = \omega_1^2(1+\eta_p^2)\prod_{i=2}^{p-1}(1+\omega_i^2)$$

آنگاه:

$$\begin{aligned}
1+\sum_{i=1}^p\eta_i^2 &= 1+\sum_{i=2}^p\eta_i^2+\eta_1^2 \\
&= (1+\eta_p^2)\prod_{i=2}^{p-1}(1+\omega_i^2) \\
&+\ \omega_1^2(1+\eta_p^2)\prod_{i=2}^{p-1}(1+\omega_i^2) \\
&= (1+\eta_p^2)\prod_{i=1}^{p-1}(1+\omega_i^2)
\end{aligned}$$

در نتیجه با جایگذاری این روابط در J_k خواهیم داشت:

$$\begin{aligned}
J_k &= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\left(\frac{1}{1+\sum\eta_i^2}\right)^{\frac{k}{2}} d\eta_p\dots d\eta_1 \\
&= \int_{-\infty}^{+\infty}\left(\frac{1}{1+\eta_p^2}\right)^{\frac{k-p+1}{2}} d\eta_p \times \\
&\qquad \prod_{i=1}^{p-1}\int_{-\infty}^{+\infty}\left(\frac{1}{1+\omega_i^2}\right)^{\frac{k-i+1}{2}} d\omega_i
\end{aligned}$$

اما طبق تبصره زیر این انتگرال‌ها متناهی‌اند اگر و تنها اگر $k>1$، $k>2$، ...، $k>p$ که اشتراک‌شان شرط $k>p$ را می‌دهد.

تبصره ۱ $\int_{-\infty}^{+\infty}\left(\frac{1}{1+x^2}\right)^t dx<\infty \iff t>\frac{1}{2}$

برهان:

$$\begin{aligned}\int_{-\infty}^{+\infty}\left(\frac{1}{1+x^2}\right)^t dx &= 2\int_0^{+\infty}\left(\frac{1}{1+x^2}\right)^t dx \\ &= \beta\left(\frac{t}{2}-1,\frac{1}{2}\right)\end{aligned}$$

که ابتدا با تغییر متغیر $x=\tan u$ و سپس $\sin^2 u = y$، نتیجه حاصل می‌شود. اما $\beta(\frac{t}{2}-1,\frac{1}{2})=\frac{\Gamma(\frac{t}{2}-1)\Gamma(\frac{1}{2})}{\Gamma(t)}$. می‌دانیم تابع $\Gamma(t)$ متناهی است اگر و تنها اگر $t>0$ باشد. بنابراین انتگرال بالا متناهی است اگر و تنها اگر $t>0$ و $t>\frac{1}{2}$ که اشتراک آنها شرط $t>\frac{1}{2}$ را نتیجه می‌دهد.

۴ مثال‌ها

مثال ۱ آزمون $\mathbf{T}^2$ به شیوه‌ی بیزی:

فرض کنید $X_1,\dots,X_N$ مستقل و هریک دارای توزیع نرمال $\mathrm{N_p}(\mu,\mathbf{\Sigma})$ باشد $(N>p)$. می‌خواهیم آزمون فرض $\mathbf{H}_0:\mu=0$ را انجام دهیم.

اگر بخواهیم این آزمون را به روش کلاسیک مثلاً روش نسبت درستنمایی انجام دهیم ملاک درستنمایی بر اساس رابطهٔ $\lambda = \frac{\max_{\Sigma} L(\circ,\Sigma)}{\max_{\mu,\Sigma}(\mu,\Sigma)}$ بدست می‌آید که برای آزمون بالا λ برابر خواهد بود با:

$$\lambda^{-\frac{۲}{N}} = ۱ + \frac{T^۲}{N+۱} \qquad (۹)$$

که در آن $T^۲ = N\bar{X}'S^{-۱}\bar{X} = (N-۱)\bar{X}'A^{-۱}\bar{X}$ و **S** ماتریس کوواریانس نمونه، به صورت $S = (N-۱)^{-۱}\sum_{\alpha=۱}^{N}(X_\alpha - \bar{X})(X_\alpha - \bar{X})'$ است [۱]. آزمون نسبت درستنمایی با ناحیه بحرانی $\lambda \leq \lambda_\circ$ تعریف می‌شود. بنابراین ناحیه بحرانی آزمون به صورت $\mathbf{T}^۲ \geq k$ (یا $\lambda \leq \lambda_\circ$) به‌دست می‌آید. آمارهٔ $\mathbf{T}^۲$، به $\mathbf{T}^۲$ هتلینگ و آزمون به آزمون $\mathbf{T}^۲$ معروف است. سوالی که اکنون پیش می‌آید این است که آیا این آزمون پذیرفتنی است؟ استاین [۱۳] به کمک ساختار نمایی ابتدا با اثبات قضیه کلی برای خانواده نمایی، ثابت کرد که آزمون $\mathbf{T}^۲$ پذیرفتنی است. خلاصه‌ای از آن به شرح زیر می‌باشد:

قضیه ۲ فرض کنید $\mathbf{Y}$ به صورت یک خانواده نمایی با چگالی

$$\begin{aligned} f(y \mid \theta) &= C(\theta)\exp\{\sum_{j=۱}^{s}\theta_j Y_j(x)\} \\ &= C(\theta)\exp\{\theta' Y\} \end{aligned}$$

نسبت به اندازه $\sigma-$ میدان m روی یک فضای اقلیدسی نمونه توزیع شده باشد و همچنین فرض کنید که Ω فضای پارامتر طبیعی این خانواده باشد. و $\Omega_\circ$ زیرمجموعه حقیقی ناتهی از Ω باشد. اگر ϕ یک آزمون برای $\mathbf{H}_\circ : \theta \in \Omega_\circ$ درمقابل $\mathbf{H}_۱ : \theta \in \Omega - \Omega_\circ$ بر اساس $\mathbf{Y}' = (Y_۱, \ldots, Y_s)$ با ناحیه پذیرش $A_\circ$ باشد که در خواص زیر صدق کند، آنگاه آزمون ϕ پذیرفتنی است [۱۳]:

۱− مجموع $A_\circ$ زیرمجموعه بسته و محدب از R^s باشد.

۲− اگر برای c حقیقی، $A_\circ \cap \{\theta'\mathbf{Y} > c\}$ تهی باشد آنگاه یک $\theta^* \in \Omega$ وجود داشته باشد به طوری که برای λی به دلخواه بزرگ داشته باشیم : $\theta^* + \lambda\Theta \in \Omega - \Omega_\circ$.

استاین پس از اثبات این قضیه برای اثبات پذیرفتنی بودن $\mathbf{T}^۲$ به این صورت عمل کرد که ابتدا با یک تبدیل متعامد بر روی X_α، $(\alpha = ۱, \ldots, N)$، $Z_\alpha = \sum_{\beta=۱}^{N} c_{\alpha\beta}X_\beta$ و استفاده از این واقعیت که $B = \sum_{\alpha=۱}^{N} Z_\alpha Z'_\alpha = \sum_{\alpha=۱}^{N} X_\alpha X'_\alpha$ عناصر متناظر با $\mathbf{Y}$ و θ در قضیه بالا را مشخص کرد و سپس با جایگذاری رابطه $B = A + N\bar{X}\bar{X}'$ در رابطه (۹)، ناحیه پذیرش آزمون را به صورت $A_\circ = \{Z_N, B \mid Z'_N B^{-۱} Z_N \leq k\}$، به دست آورد که در آن $Z_N = \sqrt{N}\bar{\mathbf{X}}$ و $\mathbf{B}$ ماتریس همیشه مثبت است. آنگاه نشان داد که $A_\circ$ محدب است. بالاخره با فرض اینکه اشتراک $A_\circ$ با نیم فضای $\theta'\mathbf{Y} > c$ تهی باشد، قسمت آخر قضیه بالا را ثابت کرد و بدین ترتیب نشان داد که $\mathbf{T}^۲$ پذیرفتنی است. اما همان طور که ملاحظه شد این روش فوق‌العاده سخت و پیچیده است. اول اینکه برای اثبات قضیه به ریاضیات پیشرفته و درک فضایی بالا نیاز است. دوم اینکه وفق دادن مسئله موردنظر در شرایط قضیه و مشخص کردن عناصر $\mathbf{Y}$ و θی قضیه در مسئله، و سپس پیدا کردن θ^* در مسئله که در قسمت آخر قضیه صدق کند کار آسان و سرراستی نیست.

از اینرو این آزمون را به شیوه بیزی انجام می‌دهیم. برای

این کار تحت فرض $\mathbf{H}_0$ و $\mathbf{H}_1$ به ترتیب قرارمی‌دهیم:

$$\begin{aligned}
\theta &= [\mu, \Sigma] = [0, (I+\eta\eta')^{-1}] \\
\frac{d\Pi_0(\eta)}{d\eta} &= |I+\eta\eta'|^{-\frac{N}{2}} \\
\theta &= [\mu, \Sigma] = [(I+\eta\eta')^{-1}\eta, (I+\eta\eta')^{-1}] \\
\frac{d\Pi_1(\eta)}{d\eta} &= |I+\eta\eta'|^{-\frac{N}{2}} \exp\{\frac{1}{2}N\eta'(I+\eta\eta')^{-1}\eta\}
\end{aligned}$$

انتگرال‌پذیری چگالی‌های $\frac{d\Pi_i(\eta)}{d\eta}$ از رابطه (۸) (چون $N > p$) و از کراندار بودن $\eta'(\mathbf{I}+\eta\eta')^{-1}\eta$ (طبق رابطه (۶) و $\eta'(\mathbf{I}+\eta\eta')^{-1}\eta \leq 1$)، حاصل می‌شود. طبق رابطه (۱) داریم:

$$\begin{aligned}
f(x \mid \theta) &\propto |\Sigma|^{-\frac{N}{2}} e^{-\frac{1}{2}N\mu'\Sigma\mu} \\
&\times \exp\{-\frac{1}{2}tr\Sigma^{-1}B + N\mu'\Sigma^{-1}\bar{x}\}
\end{aligned}$$

که در آن $B = \sum_{\alpha=1}^{N} x_\alpha x'_\alpha$ و $N\bar{\mathbf{x}} = \sum_{\alpha=1}^{N} \mathbf{x}_\alpha$. آنگاه تحت فرض $\mathbf{H}_1$ خواهیم داشت:

$$\begin{aligned}
f(x \mid \theta)\frac{d\Pi_1(\eta)}{d\eta} &\propto |I+\eta\eta'|^{\frac{N}{2}} e^{-\frac{N}{2}\eta'(I+\eta\eta')^{-1}\eta} \\
&\times \exp\{-\frac{1}{2}tr(I+\eta\eta')B + N\mu'\bar{x}\} \\
&\times \frac{d\Pi_1(\eta)}{d\eta} \\
&\propto \exp\{-\frac{1}{2}tr(I+\eta\eta')B + N\mu'\bar{x}\} \\
&\propto e^{-\frac{1}{2}trB} \\
&\times \exp\{-\frac{1}{2}trB(\eta\eta' - 2N\eta'\bar{x}B^{-1})\} \\
&\propto \exp\{-\frac{1}{2}trB + \frac{1}{2}N^2\bar{x}'B^{-1}\bar{x}\} \\
&\times \exp\{-\frac{1}{2}trB(\eta - NB^{-1}\bar{x}) \\
&\qquad (\eta - NB^{-1}\bar{x})'\}
\end{aligned}$$

همچنین تحت فرض $\mathbf{H}_0$ داریم:

$$\begin{aligned}
f(x \mid \theta)\frac{d\Pi_0(\eta)}{d\eta} &\propto |I+\eta\eta'|^{\frac{N}{2}} \\
&\times \exp\{-\frac{1}{2}tr(I+\eta\eta')B\} \\
&\times |I+\eta\eta'|^{-\frac{N}{2}} \\
&\propto \exp\{-\frac{1}{2}tr(I+\eta\eta')B\}
\end{aligned}$$

سپس طبق رابطه (۲) آزمون بیزی به صورت زیر بدست می آید:

$$\frac{\int f(x|\theta)\Pi_1(d\theta)}{\int f(x|\theta)\Pi_0(d\theta)} = \exp\{\tfrac{1}{2}N^2\bar{x}'B^{-1}\bar{x}\}$$

چون به ازای هر θ احتمال پیشامد $\{N\bar{\mathbf{x}}'\mathbf{B}^{-1}\bar{\mathbf{x}} = c\}$ صفر است، پس برای $c \geq 0$، ناحیه بحرانی به صورت $\{N\bar{\mathbf{x}}'\mathbf{B}^{-1}\bar{\mathbf{x}} > c\}$ است. بنابراین وقتی که $N > p$ باشد، آزمون بیزی یکتا است و در نتیجه پذیرفتنی است. همان‌طور که در قسمت قبل اشاره شد، پذیرفتنی بودن این آزمون توسط استاین [۱۳] به کمک ساختار نمایی اثبات شده است، که به مراتب دشوارتر و طولانی‌تر است.

مثال ۲ آزمون استقلال مجموعه‌هایی از متغیرها به شیوه‌ی بیزی:

فرض کنید یک نمونه تصادفی N تایی ازبردار p—بعدی $\mathbf{X}$ که دارای توزیع نرمال $N_p(\mu, \Sigma)$ است، داریم. اگر بردار $\mathbf{X}$ به صورت $X' = (X'^{(1)}, \ldots, X'^{(q)})$، به q زیر بردار با بعدهای p_1،...، p_q افراز شود. هدف آزمون فرض استقلال زیربردارها است. به عبارت دیگر

$$H_0 : \Sigma_{ij} = 0 \qquad i \neq j$$

یا

$$H_0 : \Sigma = \Sigma_0 = \begin{bmatrix} \Sigma_{11} & 0 & \cdots & 0 \\ 0 & \Sigma_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Sigma_{qq} \end{bmatrix}$$

که در آن Σ_{ii}، از مرتبه $p_i \times p_i$ می‌باشد. همچنین فرض می کنیم $N \geq p$.

اگر پیشین‌های Π_1 و Π_0 به‌ترتیب به صورت زیر تعریف شوند،

$$\begin{aligned} \theta &= [\mu, \Sigma] = [(I + \eta\eta')^{-1}\eta Y, (I + \eta\eta')^{-1}] \\ \theta^{(i)} &= [\mu^{(i)}, \Sigma_{ii}] \\ &= [(I_{p_i} + \eta^{(i)}\eta'^{(i)})^{-1}\eta^{(i)} Y_i, (I_{p_i} + \eta^{(i)}\eta'^{(i)})^{-1}] \\ & \qquad , \ i = 1, \cdots, q \\ \frac{d\Pi_1(\eta)}{d\eta} &= |I + \eta\eta'|^{-\frac{n}{2}} \\ \frac{d\Pi_0(\eta)}{d\eta} &= \prod_{i=1}^{q} |I_{p_i} + \eta^{(i)}\eta'^{(i)}|^{-\frac{n}{2}}, \ (n = N - 1) \end{aligned}$$

که در آن توزیع شرطی Y به شرط η و Y_i به شرط η_i به ترتیب نرمال $N(0, \frac{1+\eta'\eta}{N})$ و $N(0, \frac{1+\eta'^{(i)}\eta^{(i)}}{N})$ است. همچنین $(\eta^{(1)}, Y_1)$،...،$(\eta^{(q)}, Y_q)$ متقابلاً مستقل‌اند. آنگاه این اندازه‌ها بنابر رابطه (۸) انتگرال‌پذیرند اگر و تنها اگر $N \geq p$ باشد. بنابراین

$$\begin{aligned} \int f(x \mid \theta)\Pi_1(d\theta) &= c \int\int \exp\{-\frac{1}{2} tr(I + \eta\eta')A \\ &+ N\bar{x}'\bar{x} + (Y - \bar{x}'\eta)^2\} dY d\eta \\ &= c\, e^{-\frac{1}{2} tr(A + N\bar{x}'\bar{x})} \int e^{-\frac{1}{2}\eta' A \eta} d\eta \\ &= c\, |A|^{-\frac{1}{2}} e^{-\frac{1}{2} tr(A + N\bar{x}'\bar{x})} \end{aligned}$$

که در آن $A = \sum_{\alpha=1}^{N} x_\alpha x'_\alpha - N\bar{x}\bar{x}'$ و تساوی‌ها با استفاده از روابط (۳)، (۴) و (٦) بدست آمده‌اند. به همین ترتیب داریم:

$$\begin{aligned} \int f(x \mid \theta)\Pi_0(d\theta) &= \prod_{i=1}^{q} \int f(x^{(i)} \mid \theta^{(i)}) \Pi_0^{(i)}(d\theta^{(i)}) \\ &= c \prod_{i=1}^{q} [|A_{ii}|^{-\frac{1}{2}} \\ &\times \ \exp\{-\frac{1}{2} tr(A_{ii} + N\bar{x}'^{(i)}\bar{x}^{(i)})\}] \\ &= c \left(\prod_{i=1}^{q} |A_{ii}| \right)^{-\frac{1}{2}} \\ &\times \ \exp\{-\frac{1}{2} tr(A + N\bar{x}'\bar{x})\} \end{aligned}$$

آنگاه طبق رابطه (۲) آزمون بیزی به صورت زیر بدست می آید:

$$\frac{\int f(x \mid \theta)\Pi_1(d\theta)}{\int f(x \mid \theta)\Pi_0(d\theta)} = \left(\frac{|A|}{\prod_{i=1}^{q} |A_{ii}|}\right)^{-\frac{1}{2}} \geq c$$

که در آن $o \leq c \leq \infty$. چون به ازای هر θ احتمال پیشامد $\left\{\frac{\prod_{i=1}^{q} |A_{ii}|}{|A|} = c\right\}$ صفر است، آزمون بیزی یکتا و بنابراین پذیرفتنی است. این آزمون همان آزمون نسبت درستنمایی نیز هست [۱]. بنابراین آزمون نسبت درستنمایی وقتی که $N \geq p$ باشد پذیرفتنی است. کی‌فر و شوارتز [۶] برای اولین بار به روش بیزی نشان دادند که آزمون نسبت درستنمایی پذیرفتنی است. البته آنها مسئله را در حالت ساده‌تر یعنی وقتی که بردار $\mathbf{x}$ دارای توزیع نرمال $N_p(0, \Sigma)$ باشد، حل کرده‌اند. در این مثال اگر $q = 2$ و $p_1 = 1$ باشد، آنگاه آزمون ضریب همبستگی چندگانهٔ $\mathbf{R}^2$ حاصل می‌شود و اگر در همین حالت خاص $p = 2$ باشد، آزمون دوطرفه کلاسیک $H_0 : \rho = 0$ در مقابل $H_1 : \rho \neq 0$ بر اساس همبستگی نمونه بدست می‌آید.

مراجع

[1] Anderson, T.W. (2003), *An Introduction to Multivariate Statistical Analysis*, Wiley, New York.

[2] Birnbaum, A. (1955), Characterizations of complete classes of tests of some multiparametric hypotheses with applications to likelihood ratio tests, *Ann. Math. Statist.*, 26, 21-36.

[3] Ellison, B.F. (1962), A Classification problem in which information about alternative distributions is based on samples, *Ann. Math. Statist.*, 33, 213-223.

[4] Ghosh, M.N. (1964), On the admissibility of some tests of MANOVA, *Ann. Math. Statist.*, 35, 789-794.

[5] Karlin, S. (1957), Polya type distributions, *II*, *Ann. Math. Statist.*, 28, 281-308.

[6] Kiefer, J. and Schwartz, R. (1965), Admissible Bayes character of T^2, R^2 and othet fully invariant tests for classical multivariate normal problems, *Ann. Math. Statist.*, 36, 747-770.

[7] Lehmann, E.L.(1959), *Testing Statistical Hypotheses*, Wiley, New York.

[8] Lehmann, E.L. and Stein, C. (1948), Most powerful tests of composite hypotheses.I . Normal distributions, *Ann. Math. Statist.*, 19, 495-516.

[9] Lehmann, E.L. and Stein, C. (1953), The admissibility of certain invariant statistical tests involving a translation parameter, *Ann. Math. Statist.*, 24, 473-479.

[10] Nandi, H.K. (1963), On the admissibility of a class of tests, *Calcutta. Statist. Assoc. Bull.*, 15, 13-18.

[11] Robert, C.P. (2001), *The Bayesian Choice*, Springer, New York.

[12] Schwartz, R. (1964), Admissible invariant tests in MANOVA, *Ann. Math. Statist.*, 35, 1398.

[13] Stein, C. (1956), The admissibility of Hotelling's $T^{۲}$-test, *Ann. Math. Statist.*, 27, 616-623.

[14] Wald, A. (1942), On the power function of the analysis of variance test. *Ann. Math. Statist.*, 13, 434-439.