

SID



سرویس های ویژه



سرویس ترجمه تخصصی



کارگاه های آموزشی



بلاگ مرکز اطلاعات علمی



عضویت در خبرنامه



فیلم های آموزشی

کارگاه های آموزشی مرکز اطلاعات علمی جهاد دانشگاهی



مباحث پیشرفته یادگیری عمیق؛
شبکه های توجه گرافی
(Graph Attention Networks)



کارگاه آنلاین آموزش استفاده از
وب آو ساینس



کارگاه آنلاین مقاله روزمره انگلیسی

p -مقدار معمولی و p -مقدار اصلاح شده، چگونه بهتر قضاوت کنیم؟

حمید اسماعیلی^۱، مینا توحیدی^۲، سیدروح‌الله روزگار^۲، مهدی امیری^۲
^۱ معاونت اجتماعی و پیشگیری از وقوع جرم دادگستری کل استان بوشهر
^۲ گروه آمار، دانشگاه شیراز

تاریخ دریافت: ۱۳۸۹/۱/۲۲ تاریخ آخرین بازنگری: ۱۳۸۹/۵/۱۶

چکیده: اغلب در آزمون فرض از p -مقدار برای تصمیم‌گیری استفاده می‌شود. آیا p -مقدار بهترین معیار برای رد یا تأیید فرضیه صفر است؟ آیا می‌توان معیاری بهتر از آن در اختیار داشت؟ در این مقاله مسأله آزمون فرضیه نه به‌عنوان یک تصمیم بلکه به‌عنوان یک مسأله برآوردیابی برای احتمال رخ دادن مجموعه مشخص شده با θ در نظر گرفته می‌شود و از p -مقدار به‌عنوان برآوردگری برای احتمال رخ دادن θ استفاده خواهد شد. از طرفی در نظر گرفتن اعداد حقیقی به‌عنوان فضای پارامتری همواره مورد تأیید محققان بوده است. در حالی‌که در بسیاری از کاربردها فضای پارامتری محدود شده است. برای حالتی که فضای پارامتری کراندار باشد معیاری به نام p -مقدار اصلاح شده در توزیع نرمال برای آزمون‌های یک و دو طرفه ارائه خواهد شد که نسبت به p -مقدار معمولی عملکرد بهتری دارد.

واژه‌های کلیدی: p -مقدار اصلاح شده، قانون درست‌نمایی، فضای پارامتری محدود شده، استنباط شواهدی.

آدرس الکترونیک مسئول مقاله: حمید اسماعیلی، email: ihamid@gmail.com
کد موضوع‌بندی ریاضی (۲۰۰۰): ۶۲F۰۳

روش‌های معمول آزمون فرض یک مسأله انتخاب تصمیم بین H_0 و H_1 است، که با استفاده از p -مقدار انجام می‌شود. در این مقاله فرض می‌شود آزمون فرضیه، یک مسأله برآوردیابی در نظریه تصمیم است، که در آن برآورد کردن احتمال رخ دادن مجموعه Θ یا برآورد تابع $I_{\Theta}(\theta)$ مورد نظر است. در اینجا $I_A(\cdot)$ تابع نشانگر مجموعه A است. سؤالی که مطرح می‌شود آن است که آیا p -مقدار می‌تواند برآوردگر مناسبی برای احتمال رخ دادن مجموعه Θ باشد؟ در واقع p -مقدار به‌عنوان برآوردی برای $I_{\Theta}(\theta)$ در نظر گرفته می‌شود. آیا p -مقدار می‌تواند بهترین معیار برای رد یا تأیید فرضیه صفر باشد؟ آیا می‌توان معیاری بهتر از p -مقدار برای بهتر تصمیم گرفتن در مورد رد یا تأیید فرضیه صفر در اختیار داشت؟ انتقادات زیادی از p -مقدار توسط آماردانان بیز مطرح شده است. اخیراً دیدگاه استنباط شواهدی مطرح شده است که تنها مبتنی بر مشاهدات است و تحت تأثیر توزیع پیشینی و تابع زیان نیست و کاملاً متکی بر تابع درست‌نمایی است. این دیدگاه اولین بار توسط رویال (۱۹۹۷) مطرح شد. سپس فارستر و سابر (۲۰۰۱)، گودمن (۲۰۰۵) به بررسی و مطالعه این دیدگاه پرداختند. در ایران نیز ارقامی و همکاران (۱۳۸۷a و ۱۳۸۷b) به بررسی معیارها با استفاده از دیدگاه استنباط شواهدی پرداختند و معایب p -مقدار را از دیدگاه استنباط شواهدی مطرح کردند که در بخش ۲ مقاله به بعضی از این معایب اشاره خواهد شد. p -مقدار از دیدگاه استنباط شواهدی به‌عنوان یک معیار شواهدی خوب مطرح نیست و مورد انتقاد قرار گرفته است. اما در نظر گرفتن فضای پارامتری به صورت $(-\infty, +\infty)$ همواره مورد تردید محققان بوده است. آیا وقتی محقق میانگین قد دانش‌آموزان یک دبیرستان را بررسی می‌کند در نظر گرفتن اعداد بسیار بزرگ و نزدیک به $+\infty$ برای میانگین قد نوجوانان دبیرستانی هرچند با احتمال بسیار کم، امری عجیب نیست؟ در بسیاری از آزمون‌ها فرض می‌شود مشاهدات از توزیع نرمال با فضای پارامتری $(-\infty, +\infty)$ پیروی می‌کنند. ولی در اغلب مسائل کاربردی فضای پارامتری کراندار است. به‌عنوان مثال برای آزمون وزن نوزادان تازه متولد شده در یک بیمارستان می‌توان

باتوجه به اطلاعات قبل یا یک نمونه‌گیری اولیه یک کران بالا و پایین برای فضای پارامتری به دست آورد. مندلکرن (۲۰۰۲) مثال‌هایی ارائه کرد که وقتی فضای پارامتر کران‌دار باشد، روند کلاسیک نیمن برای بسیاری از علوم رضایت بخش نمی‌باشد. او به تناقضاتی در مورد طول بازه اطمینان برای θ وقتی $X \sim N(\theta, \sigma^2)$ و θ کراندار (مثلاً $\theta \geq 0$) و σ^2 معلوم است، یا برای λ وقتی $X \sim \text{Poisson}(\lambda)$ و λ کراندار باشد اشاره کرد. اگر فضای پارامتری کراندار باشد آیا p مقدار که فقط به مشاهده x و مقدار θ بستگی دارد و از اطلاعاتی که ما در مورد فضای پارامتری داریم استفاده نمی‌کند معیار مناسبی برای رد یا تأیید فرضیه صفر هست؟ وقتی $X \sim N(\theta, \sigma^2)$ ، آیا برای آزمون فرضیه‌های یک طرفه
$$\begin{cases} H_0 & : \theta \leq \theta_0 \\ H_1 & : \theta > \theta_0 \end{cases}$$
 و فرضیه‌های دو طرفه
$$\begin{cases} H_0 & : \theta = \theta_0 \\ H_1 & : \theta \neq \theta_0 \end{cases}$$
 اگر فضای پارامتری $\Theta_0 \cup \Theta_1$ به جای $(-\infty, +\infty)$ به صورت بازه (b, a) باشد، که در آن $-\infty \leq b \leq a \leq +\infty$ آیا p مقدار که به کران‌های a و b بستگی ندارد، بهترین معیار برای رد یا قبول فرضیه صفر هست؟

در بخش ۲ معایب p مقدار از دو دیدگاه آمار بیزی و آمار شواهدی بیان می‌شوند. در بخش ۳ تابع زیان مناسب معرفی خواهد شد. در بخش ۴ روایی تحت تابع زیان توان دوم خطا بررسی می‌شود. در بخش ۵ مقدار اصلاح شده معرفی می‌شود. در بخش ۶ p مقدار اصلاح شده و p مقدار معمولی از دیدگاه آمار بیزی مقایسه می‌گردد. در بخش ۷ بهتر بودن p مقدار اصلاح شده نسبت به p مقدار معمولی، با مینیمم کردن مجموع خطای نوع اول و خطای نوع دوم نشان داده خواهد شد.

۲ معایب p مقدار

با اینکه p مقدار در مواردی با احتمال بیز پسین برابر می‌شود، اما در آمار بیزی بیشترین انتقادات از آن انجام می‌شود. لیندلی (۱۹۵۷) نشان داد برای توزیع نرمال با میانگین θ ، آزمون فرضیه $H_0: \theta = \theta_0$ در مقابل $H_1: \theta \neq \theta_0$ ، بین استفاده از p مقدار و استفاده از احتمال بیز پسین تناقض وجود دارد. وی نشان داد ممکن است p مقدار موجب رد فرضیه H_0 شود، در حالی که احتمال بیز پسین H_0 مقداری شود

که منجر به تأیید فرضیه H_0 گردد. بعدها نشان داده شد تفاوت بین نتایج آزمون معنی داری و آزمون بیزی در فرضیه‌های دو طرفه به توزیع پیشینی و تعداد نمونه بستگی ندارد، در واقع اختلاف ذکر شده بستگی به مجموعه‌ای دارد که تحت آن بزرگترین کران پایین احتمال پسین فرض صفر به دست آورده می‌شود (لهمن، ۲۰۰۵). برگر و دلامپدی (۱۹۸۷) نشان دادند بین p -مقدار و احتمال بیز پسین در آزمون دو طرفه برای خانواده توزیع دو جمله‌ای تناقض وجود دارد و ثابت کردند اگر توزیع پیشینی، متقارن حول θ_0 یا θ_1 یا مزدوج با میانگین θ_0 یا هر توزیع با میانه θ_0 باشد، بین p -مقدار و احتمال بیز پسین H_0 ، اختلاف زیادی وجود دارد. همچنین برگر و سلکه (۱۹۸۷) نشان دادند در آزمون‌های دو طرفه توزیع نرمال، حداقل احتمال بیز پسین H_0 با توزیع پیشینی دلخواه به مقدار قابل ملاحظه‌ای از p -مقدار بیشتر است. کاسلا و برگر (۱۹۸۷) نشان دادند در آزمون‌های یک طرفه این تناقض وجود ندارد و p -مقدار به صورت حدی از برآوردگر بیز است.

در بسیاری از رشته‌های علوم انسانی تفسیرهای نادرستی از p -مقدار می‌شود و اغلب آن را احتمال خطای نوع اول می‌نامند، در حالی که احتمال خطای نوع اول را خود محقق تعیین می‌کند و به توزیع آماره آزمون تحت فرض H_0 بستگی دارد و تابعی از مشاهدات نیست.

p -مقدار از دیدگاه استنباط شواهدی نیز مورد انتقاد قرار گرفته است و به عنوان یک معیار شواهدی خوب در نظر گرفته نمی‌شود. از مهمترین انتقاداتی که از دیدگاه شواهدی به p -مقدار وارد می‌شود می‌توان به موارد زیر اشاره کرد:

الف- عدم وابستگی به فرض مقابل، به عنوان مثال اگر $X \sim N(\theta, \sigma^2)$ ، آنگاه p -مقدار برای هر دو آزمون فرضیه‌های $\begin{cases} H_0: \theta = 0 \\ H_1: \theta = 1 \end{cases}$ و $\begin{cases} H_0: \theta = 0 \\ H_1: \theta = 100 \end{cases}$ با مشاهده $x = 1/5$ برابر با $0/05$ خواهد شد، در حالی که اختلاف بین دو فرضیه زیاد است، کالبیره نبودن، عدم تقارن (بدین معنی که اگر جای دو فرضیه H_0 و H_1 را عوض کنیم انتظار داریم p -مقدار جدید، یک منهای p -مقدار قبلی باشد ولی این گونه نمی‌شود)، صعودی نبودن نسبت به H_0 (بدین معنی که اگر ناحیه بزرگتری از فضای پارامتری در فرضیه صفر قرار گیرد انتظار می‌رود p -مقدار با قاطعیت بیشتری

فرضیه صفر را تأیید کند ولی در بسیاری از موارد این گونه نمی شود. همچنین اگر $H_1 \rightarrow H_0$ (به طور مثال $H_0: \theta = 1 - \frac{1}{n}$ و $H_1: \theta = 1 + \frac{1}{n}$ وقتی $n \rightarrow \infty$) انتظار می رود p مقدار به سمت $\frac{1}{2}$ میل کند، یعنی به یک میزان از دو فرضیه حمایت کند، که این گونه نیست.

در این مقاله به بررسی p مقدار به عنوان یک معیار تصمیم گیری در مورد H_0 پرداخته خواهد شد. توجه شود که هدف رد یا تأیید فرضیه صفر با استفاده از p مقدار نیست، بلکه هدف استفاده از آن به عنوان معیاری برای برآورد احتمال رخ دادن مجموعه $\Theta_0(\theta)$ است.

۳ تابع زیان

از آنجایی که پارامتر دو مقدار دارد، پس برای هر تابع زیان دلخواه داریم:

$$L(\theta, \phi(x)) = \begin{cases} L(1, \phi(x)) & \theta \in \Theta_0 \\ L(0, \phi(x)) & \theta \notin \Theta_0 \end{cases}$$

لیندلی (۱۹۸۵) تابع زیانی را مناسب دانست که برآورد $I_{\Theta_0}(\theta)$ حاصل از آن با احتمال بیز پسین $P(\theta \in \Theta_0 | x)$ برابر باشد. یکی از این گونه توابع زیان تابع زیان توان دوم خطای

$$L_2(\theta, \phi) = (I_{\Theta_0}(\theta) - \phi(x))^2 \quad (1)$$

است، که برآوردگر بیزی حاصل از آن عبارتست از

$$\phi(x) = E(I_{\Theta_0}(\theta) | x) = P(\theta \in \Theta_0 | x)$$

یکی دیگر تابع زیان لگاریتمی است که به صورت

$$L(\theta, \phi) = \log |I_{\Theta_0}(\theta) + \phi(x) - 1|$$

تعریف می شود، اما تابع زیان قدر مطلق خطا مناسب نیست، زیرا تصمیم بیزی حاصل از آن به صورت زیر است:

$$\phi_1^\pi(x) = \begin{cases} 0 & P(\theta \in \Theta_0 | x) < \frac{1}{2} \\ 1 & o.w \end{cases}$$

۴ روایی تحت تابع زیان توان دوم خطا

در آزمون فرضیه‌های یک طرفه p -مقدار در بسیاری از موارد تحت تابع زیان توان دوم خطا روا است که نشان می‌دهد معیار مناسبی است. اما در آزمون فرضیه‌های دو طرفه p -مقدار نارواست. در عین حال هر چند p -مقدار آزمون دو طرفه برای توزیع نرمال نارواست، اما هیچ برآوردگر بیزی بهتری از آن وجود ندارد.

قضیه ۱: فرض کنید X_1, \dots, X_n نمونه‌ای تصادفی از توزیع $N(\theta, 1)$ باشد. p -مقدار برای آزمون فرضیه‌های یک طرفه $\begin{cases} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{cases}$ تحت تابع زیان توان دوم خطا روا است.

برهان بدون آنکه از کلیت مسأله کاسته شود فرض می‌شود $n = 1$.

$$\begin{aligned} p\text{-مقدار} &= P_{\theta_0}(X \geq x) \\ &= 1 - \Phi((x - \theta_0)) \end{aligned}$$

اگر برای θ توزیع پیشینی $N(\theta_0, r)$ در نظر گرفته شود، آنگاه:

$$\theta|x \sim N\left(\frac{rx + \theta_0}{r + 1}, \frac{r}{r + 1}\right)$$

در نتیجه

$$\begin{aligned} \lim_{r \rightarrow \infty} P(\theta \in \Theta_0 | X = x) &= \int_{-\infty}^{\theta_0} \frac{1}{\sqrt{r\pi}} \exp\left(-\frac{1}{r}(x - \theta_0)^2\right) d\theta \\ &= \Phi(\theta_0 - x) \\ &= 1 - \Phi(x - \theta_0) \\ &= p\text{-مقدار} \end{aligned}$$

بنابراین برآورد بیزی تعمیم‌یافته تحت تابع زیان توان دوم خطا، وقتی $r \rightarrow \infty$ برابر با p -مقدار خواهد شد، و چون ریسک بیزی p -مقدار متناهی است

$$(r_{\pi, P} = E(R_P(\theta))) = \frac{1}{p} \quad (۷۶ روبرت، ۲۰۰۱)$$

قضیه ۲: در آزمون یک طرفه $\left\{ \begin{array}{l} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{array} \right.$ ، با تابع زیان توان دوم خطا،

(a) اگر $f(x|\theta)$ تابع چگالی توزیع دو جمله‌ای $B(n, \theta)$ باشد، p -مقدار برابر با
$$P(x) = P_{\theta_0}(X \geq x) = \sum_{k=x}^n \frac{n!}{k!(n-k)!} \theta_0^k (1-\theta_0)^{n-k}$$
 و رواست.

(b) اگر $f(x|\theta)$ تابع چگالی توزیع پواسون $P(\theta)$ باشد، p -مقدار برابر با
$$P(x) = P_{\theta_0}(X \geq x) = \sum_{k=x}^{\infty} \frac{e^{-\theta_0}}{k!} \theta_0^k$$
 و رواست.

قضیه ۳: (هوانگ، ۱۹۹۲) برای آزمون فرض دو طرفه، با تابع زیان توان دوم خطا و $P(x)$ پیوسته، p -مقدار نارواست.

هرچند قضیه ۳ بیان می‌کند p -مقدار نارواست، اما در قضیه زیر نشان داده می‌شود که ممکن است با وجود ناروا بودن p -مقدار در آزمون دو طرفه، هیچ برآوردگر بیزی برتر از آن وجود نداشته باشد.

قضیه ۴: برای آزمون فرضیه‌های دو طرفه

$$\left\{ \begin{array}{l} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{array} \right.$$

براساس تک مشاهده X از چگالی $N(\theta, 1)$ و استفاده از تابع زیان توان دوم خطا، هیچ برآوردگر بیزی مناسبی نمی‌تواند از p -مقدار بهتر باشد.

برهان بدون آنکه از کلیت مسأله کاسته شود، فرض کنید $\theta_0 = 0$. برآوردگر بیزی این آزمون عبارتست از

$$\phi^\pi(x) = \frac{\pi_0 f(x|0)}{\pi_0 f(x|0) + (1-\pi_0) \int f(x|\theta)g(\theta)\mu(d\theta)}$$

که در آن تابع چگالی $N(\theta, 1)$ است و p -مقدار برابر است با

$$\begin{aligned} P(x) &= P(|X| \geq x) \\ &= 2(1 - \Phi(x)) \end{aligned}$$

سه حالت زیر را در نظر بگیرید:

حالت اول. اگر $\pi_0 = 1$ ، آنگاه $\phi^\pi(x) = 1$. اگر $\theta \rightarrow \infty$ آنگاه $\circ \rightarrow I_{\{\theta_0\}}$ و $1 \rightarrow \phi(x)$ در نتیجه $P(x) = 2(1 - \phi(x)) \rightarrow \circ$. بنابراین

$$\begin{aligned} R(\theta, P(x)) &= E_\theta (I_{\Theta_0}(\theta) - P(x))^2 \\ &= E_\theta (I_{\{\theta_0\}}(\theta) - P(x))^2 \rightarrow \circ \end{aligned}$$

اما

$$\begin{aligned} R(\theta, \phi^\pi(x)) &= E_\theta (I_{\Theta_0}(\theta) - \phi^\pi(x))^2 \\ &= E_\theta (I_{\{\theta_0\}}(\theta) - \phi^\pi(x))^2 = 1 \end{aligned}$$

در نتیجه ریسک p -مقدار کمتر از ریسک برآوردگر بیزی $\phi^\pi(x)$ است، بنابراین برآوردگر $\phi^\pi(x)$ از p -مقدار بهتر نیست.

حالت دوم. اگر $\pi_0 = \circ$ ، آنگاه $\phi^\pi(x) = \circ$ ، $R(\circ, \phi^\pi(x)) = 1$ و

$$\begin{aligned} R(\circ, P(x)) &= E[(P(x) - \circ)^2] \\ &= \text{Var}(P(x)) + E^2(P(x)^2) \end{aligned}$$

و از $P(x) \sim U(\circ, 1)$ نتیجه می شود

$$\begin{aligned} R(\circ, P(x)) &= \frac{1}{12} + \frac{1}{4} = \frac{1}{3} \\ R(\circ, \phi^\pi(x)) &= 1 > R(\circ, P(x)) = \frac{1}{3} \end{aligned}$$

بنابراین برآوردگر $\phi^\pi(x)$ نمی تواند بهتر از p -مقدار باشد.

حالت سوم. اگر $1 > \pi_0 > \circ$ ، وقتی $\theta \rightarrow \infty$ ، آنگاه $R(\theta, P(x))$ کوچکتر از $R(\theta, \phi^\pi(x))$ است. برای x های به میزان کافی بزرگ $\circ > a > |x| > P(x) > \phi^\pi(x)$

است، زیرا

$$\phi^\pi(x) \geq \frac{\pi_0 f(x|\circ)}{\pi_0 f(x|\circ) + (1 - \pi_0) f(x|\hat{\theta})} > P(x)$$

که در آن $x = \hat{\theta}$ ، برآوردگر ماکسیمم درست‌نمایی (MLE) برای θ است. برای $\theta \neq \theta_0$ ، اختلاف ریسک دو برآوردگر برابر است با

$$E_{\theta} (I_{\{\theta_0\}}(\theta) - \phi^{\pi}(x))^2 - E_{\theta} (I_{\{\theta_0\}}(\theta) - P(x))^2 = E_{\theta} (\phi^{\pi}(x)^2 - P(x)^2)$$

باتوجه به پیوستگی، یک $\varepsilon > 0$ وجود دارد که برای هر $a < |x| < a + \varepsilon$ ، آنگاه $\phi^{\pi}(x)^2 - P(x)^2 > \varepsilon$ بنابراین

$$E_{\theta} [\phi^{\pi}(x)^2] \geq \varepsilon P_{\theta} (a < |X| < a + \varepsilon) - P_{\theta} (|X| < a)$$

این کران پایین برای θ های بزرگ مثبت است، زیرا اگر $\theta \rightarrow \infty$ آنگاه

$$\frac{P_{\theta}(a < |X| < a + \varepsilon)}{P_{\theta}(|X| < a)} \rightarrow \infty$$

بنابراین اختلاف بین دو تابع ریسک برای θ های بزرگ همواره مثبت است و $\phi^{\pi}(x)$ نمی‌تواند بهتر از p -مقدار باشد.

برآوردگر $I_{\theta_0}(\theta)$ تحت تابع زیان‌های مناسب دیگر نیز برابر با برآوردگری است که با استفاده از تابع زیان توان دوم خطا به دست آمده. بنابراین نتایج به دست آمده برای هر تابع زیان مناسبی برقرار است.

۵ - مقدار اصلاح شده

فرض کنید فضای پارامتر دارای کران پایین a و کران بالای b باشد و X دارای توزیع $N(\theta, \sigma)$ است، وقتی فضای پارامتری $S = \Theta_0 \cup \Theta_1$ به صورت (a, b) باشد، برای آزمون فرضیه یک طرفه $H_0: \theta \in \Theta_0 = (a, \theta_0] = H_1: \theta \in \Theta_1 = (\theta_0, b)$ ، p -مقدار معمولی برای آزمون UMP براساس یک مشاهده x برابر با $P_{\theta_0}(X \geq x)$ و برای آزمون دو طرفه $H_0: \theta = \theta_0$ در مقابل $H_1: \theta \neq \theta_0$ در آزمون UMPU برابر $P(|X| \geq |x - \theta_0|)$ است. بنابراین p -مقدار معمولی تنها به مقدار مشاهده شده x و مقدار θ_0 بستگی دارد و نمی‌تواند از اطلاعاتی که در مورد فضای پارامتری وجود دارد استفاده کند. حال p -مقدار اصلاح شده به صورت زیر تعریف می‌شود:

آزمون یک طرفه:

$$r_{\theta_0}(x) = \frac{P_{\theta_0}(X \geq x) - \min_{\theta \in S} P_{\theta}(X \geq x)}{\max_{\theta \in S} P_{\theta}(X \geq x) - \min_{\theta \in S} P_{\theta}(X \geq x)}$$

آزمون دو طرفه:

$$r'_{\theta_0}(x) = \frac{P(|X| > |x - \theta_0|) - \min_{\theta \in (a,b)} P(|X| > |x - \theta|)}{\max_{\theta \in (a,b)} P(|X| > |x - \theta|) - \min_{\theta \in (a,b)} P(|X| > |x - \theta|)}$$

برای مشاهده x ، دامنه p -مقدار معمولی به ازای $\theta_0 \in S$ به صورت

$$(\min_{\theta \in S} P_{\theta}(X \geq x), \max_{\theta \in S} P_{\theta}(X \geq x))$$

است. برای تصمیم‌گیری در مورد فرضیه صفر، چنانچه p -مقدار از سطح اطمینان خاص α (۰/۰۵ یا ۰/۰۱) کمتر باشد، فرض صفر رد می‌شود. وقتی فضای پارامتری محدود شود استفاده از سطح آزمون مشابه با زمانی که فضای پارامتری نامحدود است، منطقی نیست. زیرا به ازای هر x ، p -مقدار معمولی بزرگتر از $\min_{\theta \in S} P_{\theta}(X \geq x)$ است. اگر $\alpha = \min_{\theta \in S} P_{\theta}(X \geq x)$ آنگاه به ازای هر مشاهده x ، فرضیه $H_0: \theta \leq \theta_0$ رد نخواهد شد، یعنی فرضیه صفر هر چقدر هم نادرست باشد، با هر نمونه‌ای همواره تأیید خواهد شد. بنابراین تبدیلی از p -مقدار معمولی ساخته می‌شود به گونه‌ای که $\min_{\theta \in S} P_{\theta}(X \geq x)$ در اندازه معیاری که بر علیه فرضیه صفر استفاده می‌شود، دخالت نداشته باشد. از طرفی با توجه به اینکه این معیار با $0 < \alpha < 1$ مقایسه می‌شود، باید دامنه این تبدیل فاصله $[0, 1]$ باشد. سپس از این قانون تصمیم‌گیری استفاده می‌شود، که فرضیه صفر رد می‌شود اگر این تبدیل از p -مقدار کمتر از سطح α باشد. بنابراین برای اینکه دامنه معیار بین $[0, 1]$ باشد، $P_{\theta_0}(X \geq x) - \min_{\theta \in S} P_{\theta}(X \geq x)$ را بر $\min_{\theta \in S} P_{\theta}(X \geq x) - \min_{\theta \in S} P_{\theta}(X \geq x)$ تقسیم می‌شود. بنابراین p -مقدار اصلاح شده $(r_{\theta_0}(x))$ به صورت یک تبدیل معقول از p -مقدار معمولی به دست می‌آید. به عنوان مثال، اگر فضای پارامتر (b, ∞) و $\theta_0 = b$ باشد، آزمون فرضیه‌های

$$\begin{cases} H_0 & : \theta = b \\ H_1 & : \theta \in (b, \infty) \end{cases}$$

را در نظر بگیرید، که کل فضای پارامتری در فرض مقابل قرار گرفته است. در این صورت فرضیه صفر برای هر x همواره رد می‌شود. اما p -مقدار معمولی برابر با $P_b(X \geq x)$ و بیشتر از صفر است، پس می‌توان α را طوری انتخاب کرد که از p -مقدار معمولی کمتر باشد و نتواند فرضیه صفر را رد کند. در حالی که p -مقدار اصلاح شده برابر

$$r_{\theta_0}(x) = \frac{P_{\theta_0}(X \geq x) - \min_{\theta \in S} P_{\theta}(X \geq x)}{\max_{\theta \in S} P_{\theta}(X \geq x) - \min_{\theta \in S} P_{\theta}(X \geq x)} = 0$$

است و با استفاده از آن فرضیه صفر به ازای هر مشاهده رد می‌شود.

۶ تقریب بیزی

در این بخش به مقایسه p -مقدار اصلاح شده $r_{\theta_0}(x)$ و p -مقدار معمولی تحت تابع زیان توان دوم خطا پرداخته می‌شود. هووانگ و همکاران (۱۹۹۲) برای ارزیابی p -مقدار پیشنهاد دادند از تابع زیان توان دوم خطای

$$L(e(x), \theta) = (e(x) - I(\theta \in \Theta_0))^2 \quad (2)$$

استفاده شود، که در آن $I(\theta \in \Theta_0)$ تابع نشانگر $I(\theta \in \Theta_0) = \begin{cases} 1 & \theta \in \Theta_0 \\ 0 & o.w \end{cases}$ و Θ_0 فضای پارامتری تحت فرض صفر است.

برای آزمون فرضیه یک طرفه نشان داده می‌شود که p -مقدار اصلاح شده یک برآوردگر بیزی برای $I(\theta \in \Theta_0)$ است، که از آن روا بودن p -مقدار اصلاح شده نتیجه می‌شود و برای آزمون فرضیه دو طرفه نشان داده می‌شود که برای همه θ هایی که عضو فضای پارامتری محدود شده هستند، میانگین توان دوم خطا (MSE) $\omega r_{\theta_0}(\bar{x})$ کوچکتر از میانگین توان دوم خطای p -مقدار معمولی است، که در آن ω مقدار ثابت متعلق به یک فاصله است.

۱.۶ مقایسه p -مقدارهای معمولی و اصلاح شده در آزمون فرضیه یک طرفه

برآوردگر بیزی $I(\theta \in \Theta_0)$ با توزیع پیشینی $\pi(\theta)$ و تابع زیان توان دوم خطا عبارتست از

$$\eta(x) = E(I(\theta \in \Theta_0) | X) = \frac{\int_{\Theta_0} \exp\{-\frac{1}{\varphi}(x-\theta)^2\} d\theta}{\int_{\Theta_0} \exp\{-\frac{1}{\varphi}(x-\theta)^2\} \pi(\theta) d\theta + \int_{\Theta_1} \exp\{-\frac{1}{\varphi}(x-\theta)^2\} \pi(\theta) d\theta} \quad (۳)$$

قضیه ۵: p -مقدار اصلاح شده $r_{\theta_0}(x)$ معادل برآوردگر بیزی $I(\theta \in \Theta_0)$ با توزیع پیشینی $\pi(\theta) = I_{\Theta_0 \cup \Theta_1}(\theta)$ است.

برهان برای فضای پارامتری $(b, +\infty)$ ناحیه Θ_0 به صورت (b, θ_0) است و برآوردگر بیزی $I(\theta \in \Theta_0)$ با توزیع پیشینی $\pi(\theta) = I_{\Theta_0 \cup \Theta_1}(\theta) = I_{(b, \infty)}(\theta)$ عبارت است از

$$\begin{aligned} \frac{\int_b^{\theta_0} \exp\{-\frac{1}{\varphi}(x-\theta)^2\} d\theta}{\int_b^{\infty} \exp\{-\frac{1}{\varphi}(x-\theta)^2\} d\theta} &= \frac{\int_{-\infty}^{x-\theta_0} f(t) dt - \int_{-\infty}^{x-b} f(t) dt}{-\int_{-\infty}^{x-b} f(t) dt} \\ &= \frac{\int_{b-x}^{\theta_0-x} f(t) dt}{\int_{b-x}^{\infty} f(t) dt} \\ &= \frac{\int_{-\infty}^{x-b} f(t) dt - \int_{-\infty}^{x-\theta_0} f(t) dt}{\int_{-\infty}^{x-b} f(t) dt} \\ &= \frac{P_b(X \leq x) - P_{\theta_0}(X \leq x)}{P_b(X \leq x)} \\ &= \frac{P_{\theta_0}(X \leq x) - P_b(X \leq x)}{1 - P_b(X \leq x)} \\ &= r_{\theta_0}(x) \end{aligned}$$

حالت‌های دیگر نیز به طور مشابه ثابت می‌شوند.

گزاره ۱: برآوردگر بیزی $\eta(x)$ برای $I(\theta \in \Theta_0)$ تحت تابع زیان توان دوم خطا روا است.

با توجه به گزاره ۱، p -مقدار اصلاح شده برآوردگری روا برای $I(\theta \in \Theta_0)$ برای یک فضای پارامتری محدود شده است، اما در قضیه زیر نشان داده می‌شود p -مقدار معمولی برآوردگری ناروا برای $I(\theta \in \Theta_0)$ است.

قضیه ۶ : اگر $X \sim N(\theta, 1)$ و فضای پارامتری دارای کران پایین b و $\Theta_0 \in [b, \theta_0)$ باشد، آنگاه p -مقدار معمولی برآوردگری ناروا برای $I(\theta \in \Theta_0)$ تحت تابع زیان توان دوم خطا است.

برهان : شرط لازم و کافی برای روا بودن برآوردگر $\eta(x)$ این است که توزیع‌های پیشین $\pi_0(\theta)$ در (b, θ_0) و $\pi_1(\theta)$ در (θ_0, ∞) وجود داشته باشند، به طوری که

$$\eta(x) = \frac{\int_b^{\theta_0} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_0(d\theta)}{\int_b^{\theta_0} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_0(d\theta) + \int_{\theta_0}^{\infty} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_1(d\theta)}$$

وقتی که $\pi = \pi_0 + \pi_1$ و $\int_{\theta_0 \cup \theta_1} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi(d\theta) < \infty$ برای نشان دادن ناروایی p -مقدار معمولی با استفاده از برهان خلف، فرض کنید $P_{\theta_0}(X \geq x)$ روا باشد، بنابراین:

$$P_{\theta_0}(X \geq x) = \frac{\int_b^{\theta_0} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_0(d\theta)}{\int_b^{\theta_0} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_0(d\theta) + \int_{\theta_0}^{\infty} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_1(d\theta)}$$

در نتیجه

$$\begin{aligned} P_{\theta_0}(X \leq x) &= 1 - P_{\theta_0}(X \geq x) \\ &= \frac{\int_{\theta_0}^{\infty} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_1(d\theta)}{\int_b^{\theta_0} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_0(d\theta) + \int_{\theta_0}^{\infty} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_1(d\theta)} \\ &= \frac{\int_{\theta_0}^{\infty} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_1(d\theta) + \int_{\theta_0}^{\infty} \exp\{-\frac{1}{2}(x-\theta)^2\} \pi_0(d\theta)}{\int_b^{\infty} \exp\{-\frac{1}{2}(x-\theta)^2\} (\pi_1 + \pi_0)(d\theta)} \end{aligned}$$

از آنجایی که دامنه توزیع پیشینی $\pi_0(\theta)$ ، (b, θ_0) است. بنابراین در نتیجه

$$\int_{\theta_0}^{\infty} \exp\left\{-\frac{1}{\tau}(x-\theta)^2\right\} \pi_0(d\theta) = 0$$

$$P_{\theta_0}(X \leq x) = \frac{\int_{\theta_0}^{\infty} \exp\left\{-\frac{1}{\tau}(x-\theta)^2\right\} \pi_1(d\theta)}{\int_b^{\infty} \exp\left\{-\frac{1}{\tau}(x-\theta)^2\right\} \pi_1(d\theta)}$$

پس

$$P_{\theta_0}(X \leq x) \int_b^{\infty} \exp\left\{-\frac{1}{\tau}(x-\theta)^2\right\} \pi(d\theta) = \int_{\theta_0}^{\infty} \exp\left\{-\frac{1}{\tau}(x-\theta)^2\right\} \pi_1(d\theta) \quad (۴)$$

با توجه به اینکه $X \sim N(\theta, 1)$ داریم

$$\begin{aligned} P_{\theta_0}(X \leq x) &= \int_{-\infty}^x \frac{1}{\sqrt{\tau\pi}} \exp\left\{-\frac{1}{\tau}(t-\theta_0)^2\right\} dt \\ &= \int_{\theta_0}^{\infty} \frac{1}{\sqrt{\tau\pi}} \exp\left\{-\frac{1}{\tau}(t-x)^2\right\} dt \end{aligned}$$

زیرا اگر $T \sim N(x, 1)$ آنگاه

$$\begin{aligned} \int_{\theta_0}^{\infty} \frac{1}{\sqrt{\tau\pi}} \exp\left\{-\frac{1}{\tau}(t-x)^2\right\} dt &= P(T > \theta_0) \\ &= P(T - x > \theta_0 - x) \\ &= P(Z > \theta_0 - x) \\ &= P(Z < x - \theta_0) \\ &= P_{\theta_0}(X - \theta_0 \leq x - \theta_0) \\ &= P_{\theta_0}(X \leq x) \end{aligned}$$

در نتیجه

$$\begin{aligned} P_{\theta_0}(X \leq x) &= \int_{\theta_0}^{\infty} \frac{1}{\sqrt{\tau\pi}} \exp\left\{-\frac{1}{\tau}(t-x)^2\right\} dt \\ &\propto \int_{\theta_0}^{\infty} \frac{1}{\sqrt{\tau\pi}} \exp\left\{-\frac{1}{\tau}(x^2 + t^2)\right\} dt \end{aligned}$$

از طرفی داریم

$$\exp\left\{\frac{x^2}{\gamma}\right\} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\gamma\pi}} \exp(tx) \exp\left\{-\frac{1}{\gamma}t^2\right\} dt$$

فرض کنید U و Y و W متغیرهای تصادفی مستقلی باشند که Y مقادیر 0 و 1 را با احتمالات $\int_{-\infty}^{\theta_0} \frac{1}{\sqrt{\gamma\pi}} \exp\left\{-\frac{1}{\gamma}t^2\right\} dt$ و $\int_{-\infty}^{\infty} \frac{1}{\sqrt{\gamma\pi}} \exp\left\{-\frac{1}{\gamma}t^2\right\} dt$ می‌گیرد. U دارای تابع چگالی $\frac{1}{\sqrt{\gamma\pi}} \exp\left\{-\frac{1}{\gamma}u^2\right\}$ و $W = YW_0 + (1 - Y)W_1$ است که W_0 و W_1 متغیرهای تصادفی با توابع چگالی $\exp\left\{-\frac{1}{\gamma}w^2\right\}\pi_0(w)$ و $\exp\left\{-\frac{1}{\gamma}w^2\right\}\pi_1(w)$ هستند. در این صورت

$$\begin{aligned} P_{\theta_0}(X \leq x) &= \int_{\theta_0}^{\infty} \frac{1}{\sqrt{\gamma\pi}} \exp(tx) \exp\left\{-\frac{1}{\gamma}(t+x)^2\right\} d\theta \\ &= \exp\left\{-\frac{x^2}{\gamma}\right\} \int_{\theta_0}^{\infty} \frac{1}{\sqrt{\gamma\pi}} \exp(tx) \exp\left\{-\frac{1}{\gamma}t^2\right\} dt \\ &= \frac{1}{\sqrt{\gamma\pi}} \exp\left\{-\frac{x^2}{\gamma}\right\} \int_{\theta_0}^{\infty} \exp(ux) \exp\left\{-\frac{1}{\gamma}u^2\right\} du \end{aligned}$$

از طرفی

$$\begin{aligned} \int_b^{\infty} \exp\left\{-\frac{1}{\gamma}(x-\theta)^2\right\} \pi(d\theta) &= \int_b^{\theta_0, \infty} \exp\left\{-\frac{1}{\gamma}x^2\right\} \exp(\theta x) \pi(d\theta) \\ &= \exp\left\{-\frac{x^2}{\gamma}\right\} \int_b^{\infty} \exp\left\{-\frac{1}{\gamma}\theta^2\right\} \exp(\theta x) \pi(d\theta) \\ &= \exp\left\{-\frac{x^2}{\gamma}\right\} \int_b^{\infty} \exp\left\{-\frac{1}{\gamma}w^2\right\} \exp(wx) \pi(dw) \\ &= \exp\left\{-\frac{x^2}{\gamma}\right\} \int_b^{\infty} \exp\left\{-\frac{1}{\gamma}w^2\right\} \exp(wx) \pi(w) dw \end{aligned}$$

پس سمت چپ رابطه (۴) برابر است با

$$\begin{aligned} P_{\theta_0}(X \leq x) &= \int_b^{\infty} \exp\left\{-\frac{1}{\gamma}(x-\theta)^2\right\} \pi(d\theta) \\ &= \exp\left(-x^2\right) \int_{\theta_0}^{\infty} \frac{1}{\sqrt{\gamma\pi}} \exp(ux) \exp\left\{-\frac{1}{\gamma}u^2\right\} du \int_b^{\infty} \exp\left\{-\frac{1}{\gamma}w^2\right\} \exp(wx) \pi(w) dw \end{aligned}$$

و سمت راست رابطه (۴) نیز برابر است با

$$\begin{aligned} \int_{\theta_0}^{\infty} \exp\left\{-\frac{1}{\gamma}(x-\theta)^2\right\} \pi(d\theta) &= \exp\left\{-\frac{x^2}{\gamma}\right\} \int_{\theta_0}^{\infty} \exp\left\{-\frac{1}{\gamma}\theta^2\right\} \exp(\theta x) \pi(d\theta) \\ &= \exp\left\{-\frac{x^2}{\gamma}\right\} \int_{\theta_0}^{\infty} \exp\left\{-\frac{1}{\gamma}w^2\right\} \exp(wx) \pi_0(w) dw \end{aligned}$$

پس ما می توانیم رابطه (۴) را به صورت زیر بنویسیم

$$\begin{aligned} & \exp(-x^2) \int_{\theta_0}^{\infty} \frac{1}{\sqrt{2}\pi} \exp(ux) \exp\{-\frac{1}{4}u^2\} du \int_b^{\infty} \exp\{-\frac{1}{4}w^2\} \exp(wx) \pi(w) dw \\ & = \exp\{-\frac{x^2}{4}\} \int_{\theta_0}^{\infty} \exp\{-\frac{1}{4}u^2\} \exp(wx) \pi_0(w) dw \end{aligned}$$

در نتیجه

$$\begin{aligned} & \int_{\theta_0}^{\infty} \frac{1}{\sqrt{2}\pi} \exp(ux) \exp\{-\frac{1}{4}u^2\} du \int_b^{\infty} \exp\{-\frac{1}{4}w^2\} \exp(wx) \pi(w) dw \\ & = \exp\{-\frac{x^2}{4}\} \int_{\theta_0}^{\infty} \exp\{-\frac{1}{4}u^2\} \exp(wx) \pi_0(w) dw \end{aligned}$$

بنابراین

$$\begin{aligned} & \int_{\theta_0}^{\infty} \exp(ux) \exp\{-\frac{1}{4}u^2\} du \int_b^{\infty} \exp\{-\frac{1}{4}w^2\} \exp(wx) \pi(w) dw \\ & = \int_{-\infty}^{\infty} \exp\{-\frac{1}{4}u^2\} du \int_{\theta_0}^{\infty} \exp\{-\frac{1}{4}w^2\} \exp(wx) \pi_0(w) dw \end{aligned}$$

با توجه به تساوی بالا و اینکه $P(Y = 1) = P(U \geq \theta_0)$, برای هر x

$$E[e^{x(U+W)} | U \geq \theta_0] = E[e^{x(U+W)} | Y = 1]$$

از آنجا که تابع مولد گشتاور توزیع‌های شرطی برای همه x ها یکسان است، پس خود توزیع‌های شرطی یکسان هستند، یعنی برای همه مقادیر z

$$P(U + W \leq z | U \geq \theta_0) = P(U + W \leq z | Y = 1)$$

اما از آنجا که همواره $U \geq \theta_0$ و $W \geq b$ ، پس اگر $z < b + \theta_0$ آنگاه $P(U + W \leq z | Y = 1) = 0$ ، ولی مقدار $P(U + W \leq z | U \geq \theta_0) = 0$ همیشه مثبت و مخالف صفر است، که این تناقض از فرض روا بودن برای p -مقدار معمولی ناشی می شود.

۲.۶ مقایسه p -مقدارهای معمولی و اصلاح شده در آزمون فرضیه دو طرفه

در قضیه ۷ فاصله در برگیرنده w به گونه‌ای در نظر گرفته شده است که کارایی $wr_{\theta_0}(\bar{x})$ بهتر از کارایی p -مقدار معمولی تحت تابع زیان توان دوم خطا باشد. در

تحلیل‌ها و روش‌های عددی، ω را می‌توان یک عدد ثابت نزدیک به یک در نظر گرفت. بنابراین p -مقدار اصلاح شده دارای کارایی بهتری نسبت به p -مقدار معمولی تحت تابع زیان توان دوم خطا است.

قضیه ۷: فرض کنید \bar{X} دارای توزیع $N(\theta, 1)$ و فضای پارامتری θ به صورت (a, b) باشد. تحت تابع زیان توان دوم خطا (۲)

(i) برای $\theta = \theta_0$ ، $\omega r_{\theta_0}(\bar{x})$ دارای میانگین توان دوم خطا کمتری نسبت به p -مقدار معمولی است وقتی که ω متعلق به بازه زیر باشد:

$$\frac{E_{\theta_0}(r_{\theta_0}(\bar{x})) \pm \sqrt{(E_{\theta_0}(r_{\theta_0}(\bar{x})))^2 - E_{\theta_0}(r_{\theta_0}(\bar{x})) (1 - E_{\theta_0}(P_{\theta_0}(\bar{x}) - 1))^2}}{E_{\theta_0}(r_{\theta_0}(\bar{x}))}$$

(ii) برای $\theta \neq \theta_0$ ، $\omega r_{\theta_0}(\bar{x})$ میانگین توان دوم خطای کمتر نسبت به p -مقدار معمولی است وقتی که ω متعلق به بازه زیر باشد:

$$\left(0, \min_{\theta \in (a,b), \theta \neq \theta_0} \left\{ \frac{E_{\theta}[P_{\theta_0}^2(\bar{X})]}{E_{\theta}[r_{\theta_0}^2(\bar{X})]} \right\} \right) \quad (5)$$

برهان: (i) وقتی $\theta = \theta_0$ ، آنگاه $I(\theta \in \Theta_0) = 1$

$$MSE(P_{\theta_0}(\bar{x})) = E_{\theta_0}[P_{\theta_0}(\bar{X}) - 1]^2$$

$$MSE(\omega r_{\theta_0}(\bar{x})) = E_{\theta_0}[(\omega r_{\theta_0}(\bar{X}) - 1)^2]$$

حال از نامساوی $MSE(\omega r_{\theta_0}(\bar{x})) < MSE(P_{\theta_0}(\bar{x}))$ نتیجه می‌شود:

$$\omega^2 E_{\theta_0}[r_{\theta_0}^2(\bar{x})] - 2\omega E_{\theta_0}[r_{\theta_0}(\bar{x})] + 1 - E_{\theta_0}(P_{\theta_0}(\bar{x}) - 1)^2 < 0$$

که برای برقراری آن، باید ω در فاصله (۱) قرار گیرد.

(ii) وقتی $\theta \neq \theta_0$ ، آنگاه $I(\theta \in \Theta_0) = 0$. پس

$$MSE(P_{\theta_0}(\bar{x})) = E_{\theta}[(P_{\theta_0}(\bar{X}))^2]$$

$$MSE(\omega r_{\theta_0}(\bar{x})) = E_{\theta_0}[(\omega r_{\theta_0}(\bar{X}))^2]$$

بنابراین

$$E_{\theta_0}[(P_{\theta_0}(\bar{X}))^2] - \omega^2 E_{\theta_0}[(r_{\theta_0}(\bar{X}))^2] > 0$$

در نتیجه

$$\omega^2 < \frac{E_{\theta_0}[(P_{\theta_0}(\bar{X}))^2]}{E_{\theta_0}[(r_{\theta_0}(\bar{X}))^2]}$$

یعنی ω متعلق به بازه (۱) است.

فرع ۱ : با توجه به قضیه ۷ اگر یک ω وجود داشته باشد که شرایط (i) و (ii) برقرار باشند، آنگاه برای هر مقدار θ از فضای پارامتری (a, b) ، $\omega r_{\theta_0}(\bar{x})$ بهتر از p مقدار معمولی است.

۷ مینیم کردن مجموع خطای نوع اول و دوم در آزمون فرض ساده

در این بخش دو p مقدار از نقطه نظر مینیم کردن مجموع خطای نوع اول و دوم در آزمون فرضیه ساده با هم مقایسه خواهند شد. برای $X \sim N(\theta, 1)$ و فرضیه‌های ساده

$$H_0 : \theta = \theta_0 \quad H_1 : \theta = \theta_1 \quad (\theta_1 > \theta_0) \quad (6)$$

اگر $x = \theta_0$ باشد، p مقدار معمولی با استفاده از ناحیه بحرانی آزمون UMP برابر است با:

$$P_{\theta_0}(X \geq \theta_0) = 0.5$$

با توزیع پیشینی $\pi(\theta) = I_{\theta_0 \cup \theta_1}(\theta)$ برآوردگر بیزی برای این آزمون با توجه به رابطه (۳) به صورت

$$\eta(x) = \frac{f_{\theta_0}(x)}{f_{\theta_0}(x) + f_{\theta_1}(x)}$$

است، که فرضیه H_0 رد می‌شود اگر $\eta(x) < k$ همچنین

$$\eta(x) = \begin{cases} \frac{1}{3} & x = \frac{\theta_0 + \theta_1}{2} \\ > \frac{1}{3} & x < \frac{\theta_0 + \theta_1}{2} \\ < \frac{1}{3} & x > \frac{\theta_0 + \theta_1}{2} \end{cases}$$

اگر $k = 0.5$ باشد، آزمون با ناحیه رد $\{x : x > \frac{\theta_0 + \theta_1}{2}\}$ با $\{x : \eta(x) < k\}$ توجه به گزاره زیر دارای ویژگی بهینه برای مینیمم کردن مجموع دو نوع خطا است. گزاره ۲: اگر متغیر تصادفی X با تابع چگالی $f_{\theta}(x)$ دارای خاصیت MLR در X باشد، در آزمون (۵) یک ناحیه رد که مجموع خطای نوع اول و دوم را مینیمم کند، به صورت $\{x : x > c^*\}$ است که c^* از رابطه $f_{\theta_0}(c^*) = f_{\theta_1}(c^*)$ به دست می آید.

برهان با توجه به لم نیمن اسپرسون ناحیه ردی که مجموع خطای نوع اول و دوم را مینیمم کند باید به صورت $S_c = \{x : x > c\}$ باشد، که در آن مقداری ثابت است. در این صورت

$$\begin{aligned} \alpha &= P_{\theta_0}(X \geq c) \\ &= \int_c^{\infty} f_{\theta_0}(x) dx \\ \beta &= P_{\theta_1}(X < c) \\ &= \int_{-\infty}^c f_{\theta_1}(x) dx \\ e(c) &= \alpha + \beta \\ &= \int_c^{\infty} f_{\theta_0}(x) dx + \int_{-\infty}^c f_{\theta_1}(x) dx \end{aligned}$$

چون

$$\frac{\partial}{\partial c} e(c) = f_{\theta_1}(c) - f_{\theta_0}(c) = 0$$

اگر جواب معادله بالا c^* در نظر گرفته شود، آنگاه مقداری است که $e(c)$ را مینیمم می کند و $f_{\theta_0}(c^*) = f_{\theta_1}(c^*)$.

برای توزیع $N(\theta, 1)$ ، $c^* = \frac{\theta_0 + \theta_1}{2}$ است. پس ناحیه ردی که مجموع خطای نوع اول و دوم را مینیمم کند به صورت $\{x : x > \frac{\theta_0 + \theta_1}{2}\}$ است، که همان ناحیه بحرانی در آزمون بیزی است. پس ناحیه رد به دست آمده از مقدار اصلاح شده، ناحیه ردی را ایجاد می کند که مجموع خطای نوع اول و دوم آزمون متناظر با آن مینیمم می شود. از طرفی ناحیه رد به دست آمده از مقدار معمولی برابر است با

$\{x : P_{\theta_0}(X \geq x) < \alpha/5\}$ ، که معادل با ناحیه $\{x : x > \theta_0\}$ است و با ناحیه ردی که مجموع خطای نوع اول و دوم را مینیمم می کند متفاوت است.

بحث و نتیجه گیری

در این مقاله ابتدا مقدار p مقدار به عنوان برآوردگری برای احتمال رخ دادن مجموعه مشخص شده با θ_0 ، و نه به عنوان عددی برای رد یا تأیید فرضیه صفر، از دیدگاه نظریه تصمیم بررسی گردید. مقدار p به طور کلی در آزمون های یک طرفه روا است و به صورت حدی از برآوردگر بیزی است. اما در حالت دو طرفه نارواست، مطابق با هیچ برآوردگر بیز تعمیم یافته ای نیست و توانایی هیچ برآوردگر بیزی از مقدار p در آزمون های دو طرفه بیشتر نیست مگر آنکه بتوان یک برآوردگر خاص برتر پیدا نمود، که نشان دهنده جایگاه مناسب مقدار p در میان برآوردگرهای دیگر است. بنابراین در آزمون های دو طرفه نیز مقدار p می تواند یک اندازه معیار مناسب برای تأیید یا رد فرض صفر باشد. در حضور وجود اطلاعاتی درباره فضای پارامتری مقدار اصلاح شده، برای فضای پارامتری کراندار معرفی شد که نسبت به مقدار معمولی عملکرد بهتری داشته و از دیدگاه آمار بیزی و مینیمم کردن مجموع خطاهای نوع اول و دوم از مقدار معمولی بهتر عمل می کند.

تقدیر و تشکر

نویسندگان از پیشنهادهای ارزنده داوران و سردبیر محترم مجله کمال سپاس گزاری را دارند.

مراجع

ارقامی، ن.، دست برآورده، ع. و زمان زاده، الف. (۱۳۸۷)، شواهد آماری، قسمت دوم: قانون درستنمایی و معایب مقدار به عنوان پشتیبانی داده از فرضیه صفر. اندیشه آماری، ۱۲، ۲، ۳-۱۶.

ارقامی، ن. ، زمانزاده، الف. و راحتی، س. (۱۳۸۷)، شواهد آماری، تفسیر نتایج
آزمون‌های آماری کلاسیک. اندیشه آماری، ۱۲، ۱، ۳-۱۰.

کاظم‌نژاد، الف.، محبی، م. و ثناگو، م. (۱۳۸۱)، p -مقدار به‌عنوان معیاری برای
پشتیبانی فرض‌ها، محاسن و معایب، مجله اندیشه آماری، ۷، ۲، ۱۹-۲۴.

Berger, J. O. and Delmpady, M. (1987), Testing Precise Hypotheses (with
Discussion), *Statistical Science*, **2**, 317-352.

Berger, J. O. and Sellke, T. (1987), Testing a Point Null Hypothesis: The
Irreconcilability of p -value and Evidence (with discussion), *Journal of
the American Statistical Association*, **82**, 112-139.

Casella, G. and Berger, R. L. (1987), Comment on Testing Precise Hy-
pothesis by Berger, J. O. and Delmpady, *Mathematical, Statistical
Science*, **2**, 344-347.

Forster, M. and Sobber, E. (2001), Why Likelihood? In *The Nature of
Scientific Evidence, Empirical, Statistical and Philosophical Consid-
eration*, University of Chcago Press.

Goodman, S. N. (2005), Introduction to Bayesian Methods I: Measuring
The Strength of Evidence, *Clinical Trials*, **2**, 282-290.

Hwang, J. T., Casella, G., Robert, C., Wells, M. T., and Farrell, R. H.
(1992), Estimation of Accuracy in Testing, *Annals of Statistics*, **20**,
490-509.

Lehmann, E. L. (2005), *Testing Statistical Hypotheses*, New York, John
Wiley.

Lindley, D. V. (1957), A Statistical Paradox, *Biometrika*, **44**, 187-192.

- Lindley, D. V. (1985), *Making Decisions*, 2nd ed, New York, John Wiley.
- Mandelkern, M. (2002), Setting Confidence Intervals for Bounded Parameters, *Statistical Science*, **17**, 149-172.
- Robert, P. C. (2001), *The Bayesian Choice*, 2nd ed. Springer-Verlag, New York.
- Royall, R. (1997), *Statistical Evidence: A Likelihood Paradigm*, Chapman & Hall /CRC.

SID



سرویس های
ویژه



سرویس ترجمه
تخصصی



کارگاه های
آموزشی



بلاگ
مرکز اطلاعات علمی



عضویت در
خبرنامه



فیلم های
آموزشی

کارگاه های آموزشی مرکز اطلاعات علمی جهاد دانشگاهی



مباحث پیشرفته یادگیری عمیق؛
شبکه های توجه گرافی
(Graph Attention Networks)



کارگاه آنلاین آموزش استفاده از
وب آوساینس



کارگاه آنلاین مقاله روزمره انگلیسی