معیارهای بهینگی برای تشخیص مدل درست از میان مدل‌های رقیب

هوشنگ طالبی [۱] ، فریبا زاده‌لباف [۲]

چکیده:
در بسیاری از مواقع، آزمایشگر با حالتی مواجه می‌شود که چند مدل رقیب وجود دارد که تنها یکی از آن‌ها درست است اما او قادر به تعیین مدل صحیح نیست. هدف از این مقاله معرفی معیارهای بهینگی برای یافتن طرح مطلوبی است که قادر به بررسی و تشخیص مدل درست باشد. در این راستا به دو رویکرد بیزی و غیربیزی می‌پردازیم. اگرچه رویکرد بیزی محدودیت‌هایی دارد، اما از این نظر که وابستگی استنباط‌های انجام شده نسبت به مقدار اولیه پارامترها را به میزان قابل توجهی کاهش می‌دهد، نسبت به رویکرد غیربیزی مزیت خواهد داشت. در این مقاله علاوه بر معرفی معیارهای گوناگون، برای هریک الگوریتمی نیز ارائه می‌دهیم که در به‌دست‌آوردن طرح‌های بهینه به‌صورت عددی سودمند واقع می‌شوند

واژه‌های کلیدی: تشخیص مدل، طرح بهینه، مدل‌های غیرخطی، $KL-$بهینگی.

۱ مقدمه

در بسیاری از زمینه‌های تحقیقاتی معمولاً آزمایش‌هایی توسط پژوهشگران برای کشف حقایقی نظیر تشخیص عوامل مؤثر بر فرآیند، محاسبه اثرات این عوامل و نیز برازش مدلی مناسب برای متغیر پاسخ (ویژگی مورد بررسی در فرآیند) به‌منظور پیش‌بینی و بهینه‌سازی فرآیند انجام می‌گیرد. درواقع آزمایش یکی از ابزارهای اساسی در به دست‌آوردن یافته‌های جدید و یا درک بهتر وقایع است. یک طرح آماری آزمایش، طرح‌ریزی یک فرآیند، تعیین مکان و تعداد نقاط آزمایشی و تعداد تکرار آن‌هاست که با استفاده از روش‌های آماری می‌توان آن‌ها را تحلیل کرد. واضح است که چگونگی انجام یک آزمایش و نحوه جمع‌آوری داده‌ها تأثیر بسزایی بر نتایج، تحلیل‌ها و اعتبار آن‌ها دارد. بنابراین پیش از اجرای هر آزمایشی، تعیین روش اجرای آن از اهمیت زیادی برخوردار است و باوجود روش‌های گوناگون، هر آزمایشگر مایل به انتخاب طرحی است که پاسخ‌گوی نیازهایش باشد. لذا با توجه به ضرورت اجرای یک آزمایش با حداکثر کارایی باید شیوه‌ای علمی در طراحی آزمایش به‌کار گرفته شود. طرح‌های بهینه محقق را قادر می‌سازند تا با انتخاب طرحی متناسب باخواسته‌هایش به اهداف خود دست یابد. بهینگی در طرح‌ها که عبارتست از انتخاب بهینه مکان نقاط آزمایشی و تعداد اجراها در هریک از این نقاط یکی از این شیوه‌های علمی است که در به‌دست آوردن طرح‌های کارا نقشی اساسی

[۱] گروه آمار، دانشگاه اصفهان
[۲] گروه آمار، دانشگاه اصفهان

دارد. درحقیقت طرح‌های بهینه رده‌ای از طرح‌های آزمایشی هستند که با توجه به معیارهای آماری گوناگون بهترین‌اند. موضوع طرح‌های بهینه برای اولین بار در سال ۱۹۱۸ توسط اسمیت[۳] مطرح شد و پس از یک وقفه زمانی محققینی ازجمله چرنوف[۴] [۳] و کیفر[۵] [۴و۵] به آن پرداختند.

در اغلب موارد آزمایشگر با چندین مدل ممکن روبروست که تنها یکی از آن‌ها مدل درست است. در این حالت ممکن است معیار بهینگی، توان طرح در تمایز بین مدل‌ها باشد.

باید توجه داشت که در مبحث تشخیص مدل معیارهای AIC و BIC با استفاده از مشاهدات به‌دست‌آمده به انتخاب مدل درست می‌پردازند. ممکن است با تغییرات اندکی در مشاهدات، این معیارها مدل دیگری را به‌عنوان مدل درست مشخص کنند. پس اگر داده‌ها در نقاط مناسبی جمع‌آوری نشده باشند، این چنین معیارهایی به نتایج گمراه کننده منجر شده و مدل را به‌درستی انتخاب نمی‌کنند. بنابراین لازم است مشاهدات به‌گونه‌ای باشند که معیارهای انتخاب مدل در برابر آن‌ها پایدار بوده و تصمیم این معیارها در تعیین مدل درست، با تغییر مشاهدات، از ثبات کافی برخوردار باشد. معیارهایی که در این مقاله معرفی می‌شوند، برای تعیین نقاطی است که پس از به‌دست‌آمدن مشاهدات در آن نقاط، تفاوت دو مدل را تا حد امکان آشکار ساخته و تشخیص بین آن‌ها را ساده‌تر کند و به این ترتیب معیارهای انتخاب مدل را در برابر مشکلات گفته شده ایمن می‌سازند.

بنابراین طرح‌های بهینه به‌دست‌آمده از معیارهای بهینگی برای تمایز بین دو مدل، در ادبیات موضوع از اهمیت بسزایی برخوردار هستند. آتکینسون و فدرف[۶] [۱] طرح‌های T—بهینه[۷] را به این منظور معرفی کردند و سایر محققین این نتایج را برای حالت‌ها و شرایط مختلف توسعه دادند. یکی از تحقیقات اساسی در این زمینه، لوپز—فیدالگو[۸] و همکاران [۶] است که با استفاده از فاصله کولبک—لیبلر معیار KL—بهینگی[۹] را برای تمایز دو مدل در شرایط خطی و غیرخطی ارائه کردند. از آن‌جا که طرح‌های T—بهینه و KL—بهینه به مدل درست فرض شده و مقادیر عددی پارامترهای آن وابسته‌اند ناگزیریم مقادیر معلوم اولیه‌ای را جایگزین پارامترهای مجهول سازیم. از این‌رو این‌گونه طرح‌ها را بهینه موضعی[۱۰] می‌نامند. برای کاهش وابستگی طرح‌های موضعی به مدل و مقادیر اولیه پارامترها، به‌جای در نظر گرفتن یک مدل به‌عنوان مدل درست، احتمال‌های پیشینی برای درستی هریک از مدل‌ها قرار می‌دهند. هم‌چنین به‌جای قرار دادن یک مقدار خاص برای پارامترها، یک توزیع احتمالی پیشین برای آن‌ها در نظر می‌گیرند. پونسه دلن[۱۱] و آتکینسون [۷] با استفاده

---

[۳] $Smith$
[۴] $Chernoff$
[۵] $Kiefer$
[۶] $AtkinsonandFedorov$
[۷] $T - optimumdesigns$
[۸] $Lopez - Fidalgo$
[۹] $KL - optimalitycriterion$
[۱۰] $locallyoptimal$
[۱۱] $PoncedeLeon$

از رویکردی بیزی معیار بهینگی T را کلیت بخشیدند. پس از آن توماسی[۱۲] و لوپز [۹] به پیروی از آن‌ها معیار بهینگی KL را به حالت بیزی گسترش دادند.
مقاله حاضر به شکل زیر تنظیم شده است.
در بخش ۲ به مرور برخی مفاهیم مورد نیاز در بخش‌های بعد خواهیم پرداخت. در بخش ۳ معیار $T-$بهینگی را معرفی و الگوریتم مربوط به آن را به‌طور مختصر شرح می‌دهیم. پس از آن، در بخش ۴ به معیار کلی‌تر KL و الگوریتم آن می‌پردازیم. دو بخش پایانی ۵ و ٦ نیز به‌ترتیب به رویکرد بیزی دو معیار T و KL و الگوریتم‌های آن‌ها اختصاص یافته است.

۲ تعاریف

زیرمجموعه $\{x_1, x_2, ..., x_N\}$ از اعضای مجموعه فشرده X، که بعضی از آن‌ها ممکن است بر هم منطبق باشند، را یک طرح آزمایشی دقیق[۱۳] (یا گسسته) با حجم N می‌نامند. با توجه به احتمال یکسان بودن بعضی از اعضا، اگر n $(\leq N)$ نقطه مجزا، هریک با r_i تکرار وجود داشته باشد، آن را با x_1، x_2، ...، x_n نمایش می‌دهیم.
برای هر x_i ضرایب وزنی را به‌صورت $w_i = r_i/N$ در نظر بگیرید، $(i = 1, ..., n)$. یک اندازه احتمال گسسته که به‌صورت زیر نمایش داده می‌شود

$$\xi = \begin{pmatrix} x_1 ... x_n \\ w_1 ... w_n \end{pmatrix}, \qquad (1)$$

را یک طرح دقیق نُرم شده[۱۴] (گسسته) یا یک طرح $n-$نقطه‌ای با حجم N گویند.
به‌طور کلی میدان سیگمایی حاصل از زیرمجموعه‌های بورل B در X را که شامل تمام زیرمجموعه‌های تک‌نقطه‌ای است، در نظر بگیرید. اندازه احتمال تعریف شده بر (X, B) یک طرح تقریبی[۱۵] (پیوسته) نامیده می‌شود. مجموعه کلیه طرح‌های تقریبی را با H نمایش می‌دهیم.
بسیاری از مواقع در عمل، دستیابی به چنین طرح‌هایی مشکل بوده و معمولاً آن‌ها را به‌عنوان تقریبی برای طرح‌های گسسته در نظر می‌گیرند. فرض کنید طرحی که شامل تعداد محدودی نقطه است را با (۱) نشان دهیم، به‌طوری‌که ضرایب w_i مقادیر مثبت دلخواه باشند، $\sum w_i = 1$. واضح است که اگر ξ_1 و ξ_2 دو اندازه طرح باشند، آن‌گاه برای $0 \leq \lambda \leq 1$، $\lambda\xi_1 + (1-\lambda)\xi_2$ نیز یک اندازه طرح خواهد بود. به‌طور کلی $\xi(dx)$ را متناظر با طرح دلخواه در نظر می‌گیریم.

۱.۲ $T-$بهینگی

مدل رگرسیونی زیر را در نظر بگیرید که در آن مشاهدات y_{ik} به‌صورت زیر

$$y_{ik} = \eta_t(x_i) + \varepsilon_{ik} (i = 1, ..., n; k = 1, ..., r_i) \qquad (2)$$

داده می‌شوند، به‌طوری که نقاط طرح x_i مقادیر معلوم و متغیرهای تصادفی ε_{ik} مستقل و دارای توزیع نرمال

[۱۲] $Tommasi$
[۱۳] $exact design$
[۱۴] $normed exact design$
[۱۵] $approximate design$

با میانگین صفر و واریانس ثابت σ^2 هستند. در مباحث نظری (اما نه در مثال‌های عددی) بدون از دست دادن کلیت σ^2 را برابر ۱ قرار می‌دهیم. تابع η_t یکی از دو تابع معلوم $\eta_1(x,\theta_1)$ و $\eta_2(x,\theta_2)$ است و θ_1 و θ_2 مجموعه‌هایی از پارامترهای مجهول در ابعاد p_1 و p_2 هستند. هدف از آزمایش تعیین یکی از این دو مدل به‌عنوان مدل درست است.

می‌دانیم برآورد کمترین مربعات پارامترها در دو مدل که در حالت کلی ممکن است غیرخطی باشند، پاسخ‌های معادلات زیر است

$$\sum_{i=1}^{n} w_i\{y_i-\eta_j(x_i,\hat{(\theta)}_j)\}^2 = \inf_{\theta_j\in\Omega_j}\sum_{i=1}^{n} w_i\{y_i-\eta_j(x_i,\theta_j)\}^2,$$

به‌طوری‌که

$$N=\sum_{i=1}^{n} r_i\, y_i=\frac{1}{r_i}\sum_{k=1}^{r_i} y_{ik}\, w_i=r_i/N$$

و $\Omega_j\in R^{p_j}$ $(j=1,2)$ مجموعه‌هایی فشرده‌اند.

در حالت خاص، فرض کنید مدل اول درست باشد، به این معنی که $\eta_t(x)=\eta_1(x,\theta_1)$ بنابراین بهتر است آزمایش به‌گونه‌ای طراحی شود که بزرگترین مقدار ممکن برای مجموع مربعات نقص برازش مدل دوم را نتیجه دهد و به‌طور معادل مقدار زیر را ماکزیمم کند

$$\Delta_2(\xi_N)=\sum_{i=1}^{n} w_i\{\eta_t(x_i)-\eta_2(x_i,\hat{(\theta)}_{t2})\}^2 \qquad (۳)$$

که در آن

$$\sum_{i=1}^{n} w_i\{\eta_t(x_i)-\eta_2(x_i,\hat{(\theta)}_{t2})\}^2 =$$

$$\inf_{\theta_2\in\Omega_2}\sum_{i=1}^{n} w_i\{\eta_t(x_i)-\eta_2(x_i,\theta_2)\}^2$$

و ξ_N به‌صورت زیر طرح مورد نظر است

$$\xi_N=\begin{pmatrix} x_1 \ldots x_n \\ w_1 \ldots w_n\end{pmatrix}$$

طرح ξ_N^* به‌طوری که

$$\Delta_2(\xi_N^*)=\sup_{\xi_N}\Delta_2(\xi_N)$$

$T-$بهینه است. دقت کنید که به علت وابستگی ξ_N^* به مقدار θ_1، این طرح یک طرح بهینه موضعی است. در حالت کلی طرح $T-$بهینه به تعیین مدل درست و مقادیر پارامترهای رگرسیونی وابسته خواهد بود.

بنابراین در حالت کلی مایل به حل مسئله بهینه‌سازی زیر هستیم

$$\Delta_2(\xi_T^*)=\sup_{\xi\in H}\Delta_2(\xi), \qquad (۴)$$

که در آن ξ اندازه طرحی است که بر ناحیه طرح X تعریف می‌شود. فرض می‌کنیم X فشرده و برای $x\in X$، $\eta_j(x,\theta)$ پیوسته باشد.

مجموعه $\Omega_2^*(\xi)$ را به‌صورت زیر در نظر بگیرید

$$\int_X\{\eta_t(x)-\eta_2(x,\theta_2^*)\}^2\xi(dx)= \inf_{\theta_2\in\Omega_2}\int_X\{\eta_t(x)-\eta_2(x,\theta_2)\}^2\xi(dx) \qquad (۵)$$

طرح بهینه ξ_T^* که در (۴) صدق می‌کند را طرح باقاعده گویند[۱۶] هرگاه برای $\xi=\xi_T^*$، مجموعه جواب (۵) دارای جواب یکتای θ_2^* است.

[۱۶] $regular design$

قضیه ۱ فرض کنید ξ_T^* یک طرح باقاعده باشد.

(i) یک شرط لازم و کافی برای آن‌که طرح ξ_T^*، T—بهینه باشد این است که در نابرابری زیر

$$\psi_{۲}(x,\xi_T^*) \leq \Delta_{۲}(\xi_T^*)$$

صدق کند، به‌طوری که

$$\psi_{۲}(x,\xi_T^*) = \{\eta_t(x) - \eta_{۲}(x,\theta_{۲}^*)\}^{۲}.$$

(ii) در نقاط طرح بهینه، $\psi_{۲}(x,\xi_T^*)$ به کران بالای خود می‌رسد.

(iii) برای هر طرح غیر بهینه ξ، یعنی طرحی که برای آن شرط $\Delta_{۲}(\xi) < \Delta_{۲}(\xi_T^*)$ برقرار است، داریم

$$\sup_{x\in X} \psi_{۲}(x,\xi) > \Delta_{۲}(\xi_T^*).$$

اثبات: به آتکینسون و فدرف [۱] رجوع شود.

۲.۲ الگوریتم

استفاده از روش‌های نظری برای ساختن طرح‌های T—بهینه تنها در شرایط خاصی امکان‌پذیر است. در غیر این صورت می‌توان فرآند تکرار زیر را به‌کار بست.

۱— فرض کنید ξ_s، طرح در گام s—ام بوده و $x_{s+۱}$ به‌گونه‌ای باشد که در شرط زیر

$$\psi_{۲}(x_{s+۱},\xi_s) = \sup_{x\in X} \psi_{۲}(x,\xi_s)$$

صدق کند.

۲— با انتخاب α_s، $۱ \leq \alpha_s \leq ۰$، طرح بعدی را به‌صورت زیر تشکیل دهید (درواقع α_s عددی است که طرح به‌دست‌آمده در گام s—ام را به گام بعد منتقل می‌کند)

$$\xi_{s+۱} = (۱-\alpha_s)\xi_s + \alpha_s \xi(x_{s+۱}),$$

که در آن $\xi(x_{s+۱})$ طرحی است که اندازه احتمال را در تک نقطه $x_{s+۱}$ متمرکز می‌کند.

بعضی از انتخاب‌های ممکن برای دنباله α_s به قرار زیر است:

(الف) هر دنباله‌ای که در شرایط زیر صدق کند

$$\alpha_s \to ۰, \sum_{s=۰}^{\infty} \alpha_s = \infty, \sum_{s=۰}^{\infty} \alpha_s^{۲} < \infty.$$

(ب) α_s که $\Delta_{۲}\{(۱-\alpha)\xi_s + \alpha\xi(x_{s+۱})\}$ را ماکزیمم کند.

۳ KL—بهینگی

فرض کنید $f_{۱}(y,x,\theta_{۱},\tau)$ و $f_{۲}(y,x,\theta_{۲},\tau)$ دو تابع چگالی رقیب باشند که در آن‌ها $\theta_i,\ i=۱,۲$، y و x مشابه بالا تعریف می‌شوند و τ پارامتر مزاحم[۱۷] است که معمولاً به واریانس مدل‌ها مربوط می‌شود. هم‌چنین $f(y,x,\tau) = f_{۱}(y,x,\theta_{۱},\tau)$ را مدل درست در نظر می‌گیریم. با توجه به این نمادها، فاصله KL بین مدل درست $f(y,x,\tau)$ و مدل رقیب $f_{۲}(y,x,\theta_{۲},\tau)$ به‌صورت زیر خواهد بود

$$I(f,f_{۲},x,\theta_{۲},\tau) =$$
$$\int f(y,x,\tau)\log\{\frac{f(y,x,\tau)}{f_{۲}(y,x,\theta_{۲},\tau)}\}dy,\ x\in X,$$

[۱۷] $nuisance parameter$

که در آن انتگرال روی فضای نمونه‌ای مشاهدات محاسبه می‌گردد.

برای سادگی فرض می‌کنیم پارامتر مزاحم τ معلوم باشد و در نتیجه تابع واریانس، بجز مقادیر پارامترهای $\theta_i,\ i=1,2$، کاملاً معلوم خواهد بود. بنابراین با حذف τ از عبارت فوق، تاب معیار $KL-$بهینگی به‌صورت زیر تعریف می‌شود

$$I_{21}(\xi)=\min_{\theta_2\in\Omega_2}\{\int_X I(f,f_2,x,\theta_2)\xi(dx)\}.$$

پس برای یک طرح تقریبی، $\int_X I(f,f_2,x,\theta_2)\xi(dx)$ با فاصله KL بین مدل‌های f و f_2 برای N مشاهده مستقل، تقریباً متناسب خواهد بود. طرحی که $I_{21}(\xi)$ را ماکزیمم می‌کند طرح $KL-$بهینه نامیده می‌شود.

می‌توان ثابت کرد که $I(f,f_2,x,\theta_2)$ با پارامتر غیرمرکزی توزیع χ برای آزمون

$$H_0:\quad f(y,x)=f_2(y,x,\theta_2),$$

$$H_1:\quad f(y,x)=f_1(y,x,\theta_1).$$

متناسب است. بنابراین با توجه به صعودی بودن تابع توان نسبت به پارامتر غیرمرکزی، مقادیر بزرگتر $I(f,f_2,x,\theta_2)$ تابع توان بزرگتری را برای آزمون نتیجه می‌دهند. می‌توان ثابت کرد

$$I_{21}(\xi)\propto\min_{\theta_2\in\Omega_2}\{E_{H_1}(R)\},$$

که در آن R آماره آزمون فوق است. بنابراین طرح $KL-$بهینه، تابع توان را در بدترین حالت ماکزیمم می‌کند.

مطابق قبل، طرحی که برای آن مسئله بهینه‌سازی

$$\Omega_2(\xi)=\{\theta_2^*:\theta_2^*(\xi)= \qquad (۶)$$

$$arg\min_{\theta_2\in\Omega_2}\{\textstyle\int_X I(f,f_2,x,\theta_2)\xi(dx)\}\}$$

دارای پاسخ یکتا باشد، طرح باقاعده نام دارد. در غیر این صورت طرح را منفرد گویند.

پس از تعاریف فوق، به ذکر قضیه زیر می‌پردازیم. ب رای مشاهده جزئیات اثبات به لوپز و همکاران [۶] مراجعه شود.

قضیه ۲ فرض کنید ξ^*_{KL} یک طرح $KL-$بهینه باقاعده باشد.

(i) یک شرط لازم و کافی برای آن‌که طرح ξ^*_{KL}، $KL-$بهینه باشد، این است که $\psi(x,\xi^*_{KL})\leq 0,\ x\in X$، که در آن

$$\psi(x,\xi)=I(f,f_2,x,\theta_2^*)-\int_X I(f,f_2,x,\theta_2^*)\xi(dx)$$

و $\hat{\theta}_2$ پاسخ یکتای مسئله بهینه‌سازی (۶) است.

(ii) تابع $\psi(x,\xi^*_{KL})$ مقدار ماکزیمم خود را در نقاط تکیه‌گاه طرح بهینه اختیار می‌کند.

۱.۳ الگوریتم

۱− برای یک طرح مفروض ξ_s قرار دهید

$$\theta_{2,s}=arg\min_{\theta_2\in\Omega_2}\{\int I(f,f_2,x,\theta_2)\xi_s(dx)\},$$

$$x_s=arg\max_{x\in X}\{I(f,f_2,x,\theta_{2,s})\}.$$

۲− سپس برای یک α_s به‌طوری‌که $0\leq\alpha_s\leq 1$ قرار دهید

$$\xi_{s+1}=(1-\alpha_s)\xi_s+\alpha_s\xi_{x_s},$$

که در آن ξ_{x_s} طرحی است که اندازه را در تک نقطه x_s متمرکز کرده است.

شرایط معمول برای دنباله $\{\alpha_s\}$ به قرار زیر است:

$$\alpha_s \to 0, \sum_{s=0}^{\infty}\alpha_s=\infty, \sum_{s=0}^{\infty}\alpha_s^2<\infty.$$

۴ T—بهینگی بیزی

همانند قبل، مدل را به‌صورت

$$E(y)=\eta_t(x), x\in X,$$

در نظر می‌گیریم که در آن مدل درست، $\eta_t(x)$، یکی از دو تابع معلوم $\eta_1(x,\theta_1)$ یا $\eta_2(x,\theta_2)$ با احتمال‌های پیشین به‌ترتیب π_1 و $\pi_2 = 1-\pi_1$ است. فرض کنید مجموعه پارامترهای θ_j با بُعد m_j دارای توزیع احتمالی پیشین $p_j(\theta_j)$ بوده و $\Omega_j\subset R^{m_j}$ فضای پارامتر مجموعه θ_j است $(j=1,2)$.

کمیت‌های زیر

$$\Delta_1(\xi,\theta_2)=\inf_{\theta_1\in\Omega_1}\int_X\{\eta_2(x,\theta_2)-\eta_1(x,\theta_1)\}^2(dx),$$

$$\Delta_2(\xi,\theta_1)=\inf_{\theta_2\in\Omega_2}\int_X\{\eta_1(x,\theta_1)-\eta_2(x,\theta_2)\}^2(dx),$$

به‌ترتیب پارامتر غیرمرکزی مدل اول است، هنگامی‌که مدل دوم درست باشد و بالعکس.

تعمیم از (۴) به یافتن $\xi_{BT}j$ می‌پردازد به‌طوری که

$$\Gamma(\xi^*_{BT})=\sup_{\xi\in H}\Gamma(\xi),$$

که در آن

$$\Gamma(\xi)=\sum\pi_j\gamma_{\bar j}(\xi)=$$

$$\pi_1E_{\theta_1}\{\Delta_2(\xi,\theta_1)\}+\pi_2E_{\theta_2}\{\Delta_1(\xi,\theta_2)\},$$

به‌طوری که $\gamma_2=\gamma_{\bar 1}$ و $\gamma_1=\gamma_{\bar 2}$.

مطابق قبل، $\Omega_j^*(\xi)$ مجموعه جواب معادله زیر

$$\int_X\{\eta_{\bar j}(x,\theta_{\bar j})-\eta_j(x,\theta_j^*)\}^2\xi(dx)= \qquad (۷)$$

$$\inf_{\theta_j\in\Omega_j}\int_X\{\eta_{\bar j}(x,\theta_{\bar j})-\eta_j(x,\theta_j)\}^2\xi(dx)$$

است.

با فرض $\xi=\xi^*_{BT}$، اگر (۷) دارای پاسخ یکتای θ_1^* برای کلیه مقادیر $\theta_2\in\Omega_2$ و θ_2^* برای کلیه مقادیر $\theta_1\in\Omega_1$ باشد، قضیه زیر را بیان می‌کنیم.

قضیه ۳ (i) یک شرط لازم و کافی برای این‌که $\xi=\xi^*_{BT}$ T—بهینه بیزی باشد، این است که برای کلیه $x\in X$، $\psi(x,\xi^*_{BT})\le\Gamma(\xi^*_{BT})$ باشد، که در آن

$$\psi(x,\xi^*_{BT})=\sum\pi_jE_{\theta_j}\{\eta_j(x,\theta_j)-\eta_{\bar j}(x,\theta^*_{\bar j})\}^2.$$

(ii) در نقاط طرح T—بهینه بیزی، $\psi(x,\xi^*_{BT})$ کران بالای خود را اختیار می‌کند.

(iii) ربرای هر طرح غیر بهینه ξ، یعنی طرحی که برای آن داشته باشیم $\Gamma(\xi)<\Gamma(\xi^*_{BT})$، می‌توان نوشت

$$\sup_{x\in X}\psi(x,\xi)>\Gamma(\xi^*_{BT}).$$

اثبات: به پونسه دلن و آتکینسون [۷] رجوع شود.

۱.۴ الگوریتم

۱— فرض کنید ξ_s، طرح در گام s—ام بوده و x_{s+1} به‌گونه‌ای باشد که در شرط

$$\psi(x_{s+1},\xi_s)=\sup_{x\in X}\psi(x,\xi_s)$$

صدق کند.

۲− طرح بعدی را به‌صورت زیر انتخاب کنید

$$\xi_{s+1} = (1-\alpha_s)\xi_s + \alpha_s \xi(x_{s+1}).$$

همانند قبل، شرایط معمول برای دنباله $\{\alpha_s\}$ به قرار زیر است:

$$\alpha_s \to 0, \sum_{s=0}^{\infty} \alpha_s = \infty, \sum_{s=0}^{\infty} \alpha_s^2 < \infty.$$

مثال ۱ این مسئله شامل طراحی آزمایش در بازه $X = [-1, 1]$ برای تشخیص بین دو مدل رگرسیونی زیر است

$$\eta_1(x) = \theta_{10} + \theta_{11}e^{x} + \theta_{12}e^{-x},$$

$$\eta_2(x) = \theta_{20} + \theta_{21}x + \theta_{22}x^2.$$

در این‌جا مدل درست را نامعلوم فرض می‌کنیم، اما احتمال‌های پیشین برای درست بودن هر مدل در دسترس است. علاوه بر این احتمال‌ها، توزیع‌های احتمالی پیشینی که برای هر دو مدل گسسته فرض شده‌اند نیز برای پارامترهای هر یک از دو مدل موجود است. این احتمال‌ها و توزیع‌های پیشین در جدول ۱ نشان داده شده‌اند.

طرح T−بهینه بیزی که با استفاده از روش‌های متداول بهینه‌سازی به‌دست می‌آوریم، دارای پنج نقطه است و مقدار ماکزیمم تابع معیار $\Gamma(\xi)$ برای این طرح با مقدار زیر برابر است

$$\xi^*_{BT} = \begin{pmatrix} -1 & -0.681 & 0.105 & 0.847 & 1 \\ 0.244 & 0.426 & 0.254 & 0.067 & 0.009 \end{pmatrix},$$

$$\Gamma(\xi^*_{BT}) = 6.346 * 10^{-4}.$$

برای اطمینان از این‌که این طرح بهینه است، نمودار تابع $\psi(x, \xi^*_{BT})$ در شکل ۱ نشان داده شده است. درواقع مقدار ماکزیمم که در هر پنج نقطه طرحی حاصل شده است، تقریباً با $\Gamma(\xi^*_{BT})$ برابر است.

۵ KL−بهینگی بیزی

اگر $\bar{i}$ به مفهوم غیر از i و $i = 1, 2$ باشد، کمیت زیر

$$I_{\bar{i},i}(\xi, \theta_i) = \min_{\theta_{\bar{i}} \in \Omega_{\bar{i}}} \int_X l[f_i(y,x,\theta_i), f_{\bar{i}}(y,x,theta_{\bar{i}})]\xi(dx) \qquad (8)$$

هنگامی‌که $f_i(y,x,\theta_i)$ را مدل درست فرض کرده باشیم تابع معیار KL برای مدل $f_{\bar{i}}(y,x,theta_{\bar{i}})$ است.

با میانگین‌گیری روی دو مدل و بر توزیع‌های پیشین پارامترها، معیار KL−بهینگی بیزی به‌صورت زیر به‌دست می‌آید

$$I^B(\xi) = \pi_1 E_{\theta_1}[I_{21}(\xi,\theta_1)] + \pi_2 E_{\theta_2}[I_{12}(\xi,\theta_2)],$$

که در آن E_{θ_i} بیانگر امید ریاضی برحسب $p_i(\theta_i)$، یعنی توزیع احتمالی پیشین پارامتر θ_i و π_i احتمال پیشین مدل آماری $f_i(y,x,\theta_i)$، $i = 1, 2$ است. طرح ξ^*_{BKL} به‌طوری که برای آن داشته باشیم

$$\xi^*_{BKL} = arg \max_{\xi \in H} I^B(\xi),$$

طرح KL−بهینه بیزی نام دارد.

اگر فرض بر این باشد که مدل $f_i(y,x,\theta_i)$ درست است، اما تنها اطلاعات موجود درباره پارامتر θ_i توزیع پیشین $p_i(\theta_i)$ باشد، معیار KL−بهینگی جزئاً بیزی[18] به‌صورت زیر تعریف می‌شود

$$I_i^{PB}(\xi) = E_{\theta_i}[I_{\bar{i},i}(\xi,\theta_i)], i = 1, 2. \qquad (9)$$

[18] $partiallyBayesianKL - optimalitycriterion$

طرحی که $I_i^{PB}(\xi)$ را ماکزیمم می‌کند، طرح KL—بهینه جزئاً بیزی نامیده شده و با $\xi^*_{PB_i}$ نمایش می‌دهند.
توابع معیار $I_{۲۱}(\xi,\theta_۱)$ و $I_{۱۲}(\xi,\theta_۲)$ بر پایه فاصله‌های کولبک—لیبلر و متناظر با درستی به‌ترتیب مدل‌های $f_۱(y,x,\theta_۱)$ و $f_۲(y,x,\theta_۲)$ به‌دست می‌آیند. این دو تابع معیار ممکن است از لحاظ اندازه به‌قدری متفاوت باشند که طرح KL—بهینه بیزی نزدیک و یا مطابق با طرح KL—بهینه متناظر با بزرگ‌ترین مقدار تابع معیار KL باشد. برای غلبه بر این مشکل تابع معیار بیزی استانداردشده زیر مورد استفاده قرار می‌گیرد

$$\begin{aligned} I^{SB}(\xi) &= \sum_{i=۱}^{۲}\pi_i E_{\theta_i}[\frac{I_{\bar{i}i}(\xi,\theta_i)}{I_{\bar{i}i}(\xi^*_{\bar{i}i},\theta_i)}] \\ &= \sum_{i=۱}^{۲}\pi_i E_{\theta_i}[Eff_{\bar{i}i}(\xi,\theta_i)], \quad i=۱,۲, \end{aligned}$$

که در آن نسبت

$$Eff_{\bar{i}i}(\xi,\theta_i) = I_{\bar{i}i}(\xi,\theta_i)/I_{\bar{i}i}(\xi^*_{\bar{i}i},\theta_i)$$

اندازه‌ای برای کارایی طرح ξ نسبت به طرح KL—بهینه $\xi^*_{\bar{i}i}$ است. طرح KL—بهینه بیزی استانداردشده[۱۹] طرحی است که $I^{SB}(\xi)$ را ماکزیمم می‌کند. این طرح با نماد ξ^*_{SB} نمایش داده می‌شود.

طرح ξ به‌طوری که برای آن مجموعه‌های زیر

$$\Omega^*_{\bar{i}}(\xi,\theta_i) = \qquad (۱۰)$$

$$arg\min_{\theta_{\bar{i}}\in\Omega_{\bar{i}}}\int_X l[f_i(y,x,\theta_i), f_{\bar{i}}(y,x,\theta_{\bar{i}})]\xi(dx),$$

برای هر $\theta_i \in \Omega_i,\ i=۱,۲$ منحصر به‌فرد باشند، یک طرح باقاعده و در غیر این صورت منفرد نامیده می‌شود.
اگر ξ یک طرح باقاعده باشد آن‌گاه تابع زیر

$$\psi_{\bar{i}i}(x,\xi,\theta_i) = l[f_i(y,x,\theta_i), f_{\bar{i}}(y,x,\theta^*_{\bar{i}})]$$

$$-\int_X l[f_i(y,x,\theta_i), f_{\bar{i}}(y,x,\theta^*_{\bar{i}})]\xi(dx)$$

مشتق سویی $I_{\bar{i}i}(\xi,\theta_i)$ در ξ و در جهت $\delta_{\xi_x} = \xi_x - \xi$ بوده و ξ_x طرحی است که تمتم جرم را در نقطه x متمرکز کرده است و $\theta^*_{\bar{i}}$ عضو یکتای $\Omega^*_{\bar{i}}(\xi,\theta_i)$ است.
پس از محاسبات ساده، مشتق سویی $I^B(\xi,\theta_i)$ در ξ و در جهت $\delta_{\bar{\xi}} = \bar{\xi} - \xi$ به‌صورت زیر به‌دست مآید

$$\partial I^B(\xi,\bar{\xi}) = \int_X \psi^B(x,\xi)\bar{\xi}(dx), \qquad (۱۱)$$

که در آن

$$\psi^B(x,\xi) = \sum_{i=۱}^{۲}\pi_i E_{\theta_i}[\psi_{\bar{i}i}(x,\xi,\theta_i)]$$

مشتق سویی $I^B(\xi)$ در ξ و در جهت δ_{ξ_x} است.
در زیر قضیه هم‌ارزی را برای این‌گونه طرح‌ها بیان می‌کنیم. برای مشاهده اثبات، توماسی و لوپز [۹] مشاهده شود.

قضیه ۴ فرض کنید ξ^*_{BKL} یک طرح باقاعده باشد.
(i) طرح ξ^*_{BKL}، KL—بهینه بیزی است اگر و تنها اگر

$$\psi^B(x,\xi^*_{BKL}) \le ۰,\ x\in X.$$

(ii) تابع $\psi^B(x,\xi^*_{BKL})$ مقدار ماکزیمم خود را در نقاط تکیه‌گاه طرح بهینه اختیار می‌کند.

۱.۵ الگوریتم

ساختار تحلیلی طرح‌های KL—بهینه موضعی و بیزی پیچیده است. بنابراین در عمل ناگزیر به استفاده از

[۱۹] $standardized Bayesian KL - optimum design$

روش‌های عددی هستیم. در این‌جا یکی از معروف‌ترین الگوریتم‌های موجود را برای محاسبه طرح‌های KL—بهینه بیزی شرح می‌دهیم. شایان ذکر است که با تعمیمی ساده می‌توان از این روش برای یافتن طرح‌های KL—بهینه بیزی استانداردشده نیز استفاده کرد.

۱— فرض کنید که ξ_s طرح به‌دست آمده در گام s—ام باشد. برای هر نقطه از تکیه‌گاه $(i=۱,۲)$, $p_i(\theta_i)$، $\theta_{\bar{i},s}$ را به‌صورت زیر بیابید

$$\theta_{\bar{i},s} = arg \min_{\theta_{\bar{i}} \in \Omega_{\bar{i}}} \int_X l[f_i(y,x,\theta_i), f_{\bar{i}}(y,x,\theta_{\bar{i}})]\xi_s(dx)$$

و سپس مقدار زیر را محاسبه کنید

$$x_s = arg \max_{x\in X} \psi^B(x,\xi_s)$$

۲— $\circ \le \alpha_s \le ۱$ را به‌گونه‌ای انتخاب کنید که $\alpha_s \to \circ$، $\sum_{s=\circ}^{\infty} \alpha_s = \infty$، $\sum_{s=\circ}^{\infty} \alpha_s^۲ < \infty$. پس از آن $\xi_{s+۱}$ را تشکیل دهید، به‌طوری که

$$\xi_{s+۱} = (۱-\alpha_s)\xi_s + \alpha_s \xi_{x_s}$$

که در آن ξ_{x_s} طرحی است که کل جرم را در نقطه آزمایشی x_s متمرکز کرده است.

به‌سادگی می‌توان ثابت کرد که یک کران پایین برای کارایی طرح KL—بهینه بیزی، که با $Eff_{BKL}(\xi)$ نمایش داده می‌شود، به‌صورت زیر به‌دست می‌آید

$$I(\xi) = [۱ + \frac{\max_{x\in X}\psi^B(x,\xi)}{I^B(\xi)}]^{-۱} \le Eff_{BKL}(\xi) \le ۱.$$

بنابراین فرایند تکرار در گام s—ام متوقف می‌گردد، اگر برای مقدار مناسب δ مثل $\delta = ۰.۹۹$، داشته باشیم $I(\xi) > \delta$.

مثال ۲ بسیاری از مواقع در تحلیل‌های قابلیت اعتماد ممکن است توزیع‌های وایبل یا گاما را به مجموعه یکسانی از داده‌ها برازش داده و به نتیجه مطلوب رسید. هر دو توزیع معمولاً برای داده‌های چوله استفاده می‌شوند. علاوه بر این در مباحث قابلیت اعتماد و کنترل کیفیت معمولاً اطلاعات پیشین نیز در دسترس بوده و مورد استفاده قرار می‌گیرند. اگر بتوان این اطلاعات را در قالب یک توزیع احتمالی پیشین درآورد می‌توان از رویکردی بیزی در تحلیل‌ها استفاده نمود. در این قسمت به بررسی طرح KL—بهینه بیزی برای مقایسه تابع چگالی احتمال وایبل،

$$f_W y; b, c = \frac{c y^{c-۱}}{b^c} \exp[-(\frac{y}{b})^c],\ b > \circ, c > \circ$$

و تابع چگالی احتمال گاما،

$$f_G(y;\beta,\alpha) = \frac{y^{\alpha-۱}\exp(-\frac{y}{\beta})}{\beta^\alpha \Gamma(\alpha)},\ \beta > \circ, \alpha > \circ$$

می‌پردازیم.

در این جامدل‌سازی $b = \exp(\lambda_۱ x)$ و $\beta = \exp(\gamma_۱ x)$ مورد توجه خواهد بود که در آن‌ها x یک متغیر توضیحی است که در ناحیه آزمایشی $X = [۱,۲]$ تغییر می‌کند. واضح است که برای $\lambda_۱, \gamma_۱ \in R$ شروط $b > \circ$ و $\beta > \circ$ برقرار خواهند بود. با توجه به این پارامترسازی داریم

$$\begin{aligned} l[f_G(y,x,\gamma_۱,\alpha), f_W(y,x,\lambda_۱,c)] = & \quad (۱۲) \\ -\log[\Gamma(\alpha)] - \log c + (\alpha - c)\tfrac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \alpha & \\ + \tfrac{\Gamma(\alpha+c)}{\Gamma(\alpha)} \exp[c(\gamma_۱ - \lambda_۱)x] - c(\gamma_۱ - \lambda_۱)x & \end{aligned}$$

$$l[f_W(y,x,\lambda_۱,c), f_G(y,x,\gamma_۱,\alpha)] = \quad (۱۳)$$

$$\log[\Gamma(\alpha)] + \log c - ۱.۵۷۷۲ + ۰.۵۷۷۲\frac{\alpha}{c}$$
$$+\frac{\Gamma(c+۱)}{c}\exp[(\lambda_۱ - \gamma_۱)x] - \alpha(\lambda_۱ - \gamma_۱)x \quad (۱۴)$$

که به‌ترتیب فاصله‌های کولبک–لیبلر بین توابع چگالی گاما و وایبل و بالعکس خواهند بود.

اگر Y متغیری تصادفی با توزیع وایبل باشد، آن‌گاه $X = \log(Y/b)$ دارای توزیع گامبل با پارامترهای $c^{-۱}$ و ۰ خواهد بود. امید ریاضی این توزیع، شناخته شده و در این‌جا برابر است با $E_W(x) = -\gamma/c$ که در آن $\gamma \approx ۰.۵۷۷۲$ عدد اویلر[۲۰] است. بنابراین $E_W(\log Y) = \log b - \gamma/c$.

فرض کنید توزیع گاما درست باشد آن‌گاه با توجه به (۱۲) داریم

$$I_{۲۱}(\xi, \gamma_۱, \alpha) =$$
$$\min_{\lambda_۱, c}\int_X l[f_G(y,x,\gamma,\alpha), f_W(y,x,\lambda_۱,c)]\xi(dx) =$$
$$\min_{\bar\lambda_۱, c}\{-\log[\Gamma(\alpha)] - \log c + (\alpha - c)\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \alpha$$
$$+\int_X \frac{\Gamma(\alpha+c)}{\Gamma(\alpha)}\exp[c\bar\lambda_۱ x] - c\bar\lambda_۱ x\xi(dx)\},$$

که در آن $\bar\lambda_۱ = \gamma_۱ - \lambda_۱$. پس درحقیقت $I_{۲۱}(\xi,\gamma_۱,\alpha) = I_{۲۱}(\xi,\alpha)$ به $\gamma_۱$ وابسته نیست. به طریق مشابه، با فرض درستی توزیع وایبل، $I_{۱۲}(\xi,\lambda_۱,c) = I_{۱۲}(\xi,c)$ مستقل از $\lambda_۱$ است، زیرا با استفاده از (۱۳) داریم

$$I_{۱۲}(\xi, \lambda_۱, c) =$$
$$\min_{\gamma_۱,\alpha}\int_X l[f_W(y,x,\lambda_۱,c), f_G(y,x,\gamma_۱,\alpha)]\xi(dx) =$$
$$\min_{\bar\gamma_۱,\alpha}\{\log[\Gamma(\alpha)] + \log c - ۱.۵۷۷۲\frac{\alpha}{c}$$
$$+\int_X \Gamma(\frac{c+۱}{c})\exp[\bar\gamma_۱ x] - \alpha\bar\gamma_۱ x\xi(dx)\},$$

به‌طوری که $\bar\gamma_۱ = \lambda_۱ - \gamma_۱$.

همان‌گونه که پیش از این نیز بیان گردید، طرح KL–بهینه بیزی به احتمالات پیشین درست بودن هر مدل و توزیع‌های احتمالی پیشین پارامترهای دو مدل آماری که در این‌جا گسسته فرض می‌شوند، وابسته است. در این مثال، $\pi_۱ = ۰.۴$ و $\pi_۲ = ۱ - \pi_۱ = ۰.۶$ به‌ترتیب بیانگر احتمالات پیشین برای چگالی‌های و وایبل و $\theta_۱ = (\alpha, \gamma_۱)$ و $\theta_۲ = (c, \lambda_۱)$ به‌ترتیب متغیرهای تصادفی مستقل با توزیع‌های احتمالی پیشین $p_۱$ و $p_۲$ هستند. در حالات الف و ب از جدول ۲، توزیع‌های کناری پارامترهای α و c یکسان فرض شده‌اند، در حالی‌که توزیع‌های کناری $\gamma_۱$ و $\lambda_۱$ متفاوت هستند. با توجه به استقلال معیارهای KL–بهینگی از $\gamma_۱$ و $\lambda_۱$، طرح KL–بهینه بیزی برای هر دو حالت برابر بوده و به‌صورت زیر است

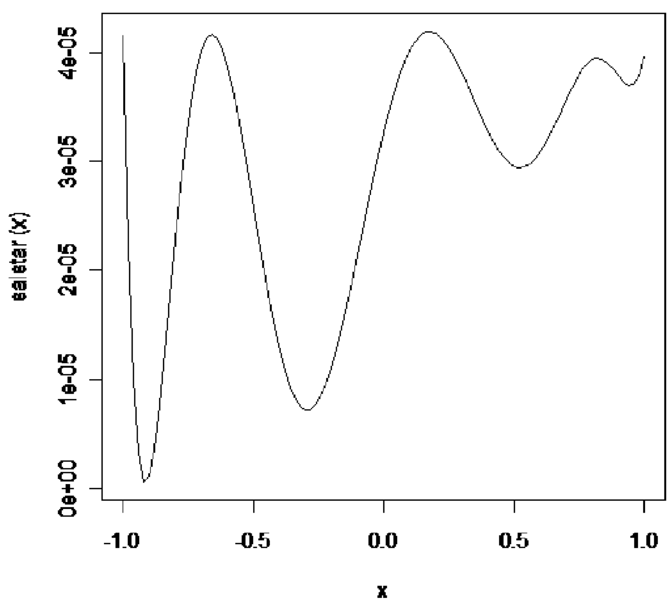
$$\xi^*_{BKL} = \begin{pmatrix} ۱ & ۲ \\ ۰.۶۱۷ & ۰.۳۸۳ \end{pmatrix}.$$

تابع $\psi^B(x, \xi^*_{BKL})$ برای این طرح نیز در شکل زیر نشان داده شده است. واضح است که این تابع مقدار ماکزیمم خود را در نقطه طرحی اختیار می‌کند.

[۲۰] $Euler number$

جدول ۱. احتمال‌های پیشین برای درستی مدل و توزیع‌های پیشین برای پارامترهای θ_1 و θ_2 در مثال ۱

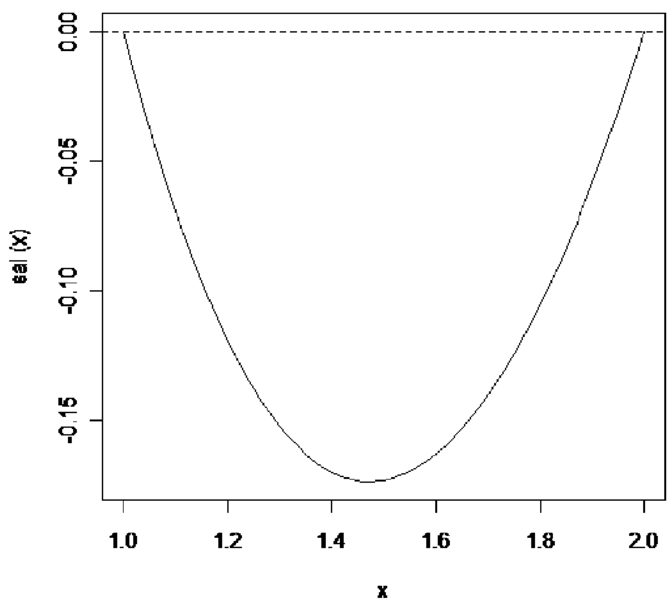
$\pi_1 = 0.6$				$\pi_2 = 0.4$			
θ_{10}	θ_{11}	θ_{12}	$p_1(\theta_1)$	θ_{20}	θ_{21}	θ_{22}	$p_2(\theta_2)$
۴.۵	−۱.۵	−۲.۰	۰.۲۵	۱.۰	۰.۵	−۲.۰	۰.۲۳
۴.۰	−۱.۰	−۲.۰	۰.۱۴	۰.۸	۰.۴	−۲.۰	۰.۳۳
۴.۵	−۲.۰	−۱.۵	۰.۱۱	۱.۰	۰.۶	−۱.۵	۰.۱۷
۵.۰	−۱.۵	−۱.۵	۰.۰۶	۱.۲	۰.۵	−۱.۵	۰.۱۵
۴.۰	−۲.۰	−۱.۰	۰.۰۵	۰.۸	۰.۶	−۱.۵	۰.۱۲
۴.۵	−۱.۵	−۱.۵	۰.۰۸				
۴.۰	−۱.۵	−۲.۰	۰.۰۵				
۴.۰	−۲.۰	−۲.۰	۰.۱۲				
۴.۵	−۲.۰	−۲.۰	۰.۰۷				
۵.۰	−۱.۵	−۲.۰	۰.۰۷				



شکل ۲. تابع $\psi^B(x, \xi^*_{BT})$ در مثال ۱.

جدول ۲. توزیع‌های احتمالی پیشین برای پارامترهای $\theta_۱ = (\alpha, \gamma_۱)$ و $\theta_۲ = (c, \lambda_۱)$ در مثال ۲

حالت الف						حالت ب					
α	$\gamma_۱$	$p_۱$	c	$\lambda_۱$	$p_۲$	α	$\gamma_۱$	$p_۱$	c	$\lambda_۱$	$p_۲$
۳	۱	۰.۱۲۵	۳	۱	۰.۱۲۵	۳	۱	۰.۲۵	۳	۱	۰.۲۵
۳	۲	۰.۱۲۵	۳	۲	۰.۱۲۵	۳	۴	۰.۲۵	۳	۴	۰.۲۵
۳	۳	۰.۱۲۵	۳	۳	۰.۱۲۵	٦	۱	۰.۲۵	٦	۱	۰.۲۵
۳	۴	۰.۱۲۵	۳	۴	۰.۱۲۵	٦	۴	۰.۲۵	٦	۴	۰.۲۵
٦	۱	۰.۱۲۵	٦	۱	۰.۱۲۵						
٦	۲	۰.۱۲۵	٦	۲	۰.۱۲۵						
٦	۳	۰.۱۲۵	٦	۳	۰.۱۲۵						
٦	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]						



شکل ۲. تابع $\psi^B(x, \xi^*_{BKL})$ در مثال ۲.

مراجع

[1] Atkinson, A.C. and Fedorov, V.V. (1975a). The design of experiments for discriminating between two rival models. *Biometrica* 62, 57-70.

[2] Atkinson, A.C. and Fedorov, V.V. (1975b). Optimal design: experiments for discriminating between several models. *Biometrica* 62, 289-303.

[3] Chernoff, H. (1953). Locally optimal design for estimating parameters. *Ann. Math. Statist.* 24, 586-602.

[4] Kiefer, J. (1958). On the nonrandomized optimality and randomized non-optimality of symmetrical design. *Ann. Statist.* 2, 849-879.

[5] Kiefer, J. (1959). Optimum experimental designs. *J. R. Statist. Soc.* B 21, 272-319.

[6] Lopez-Fidalgo, J., Tommasi, C., Trandafir, P.C. (2007). An optimal experimental design criterion for discriminating between non-normal models. *J. R. Statist. Soc.* B 69, 231-242.

[7] Ponce de Leon, A.C. and Atkinson, A.C. (1991a). Optimal experimental design for discriminating between two rival models in the presence of prior information. *Biometrica* 78, 601-608.

[8] Smith, K. (1918). On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and guidance they give towards a proper choice of the distribution of observations. *Biometrica* 12, 1-58.

[9] Tommasi, C. and Lopez-Fidalgo, J. (2010). Bayesian optimum designs for discriminating between models with any distribution. *Computational Statistics and Data Analysis* 54, 143-150.