فاصله شواهدی جانشینی برای فواصل اطمینان و اعتبار

مهدی عمادی [۱] * مرتضی محمدی [۲]
[۱] گروه آمار، دانشگاه فردوسی مشهد
[۲] گروه آمار، دانشگاه زابل، زابل، ایران

چکیده: یکی از روش‌های معتبر در بین آماردانان برای تبیین و تفسیر جوامع، استفاده از تخمین فاصله ای برای پارامترهای مورد بررسی است. از این رو در این مقاله ضمن معرفی روش‌های بدست آوردن تخمین فاصله ای، آنها را توسط چند مثال مورد مقایسه قرار می‌دهیم. برای معرفی روش‌های مرسوم می‌توان به روش بیزی، روش کلاسیک و روش نوین شواهدی اشاره کرد. واضح است که در هر روش مفروضات مخصوص خود را باید در نظر گرفت. مثلا در روش بیزی فرض بر این است که پارامتر مورد بررسی تصادفی بوده و دارای تابع احتمال است. در روش کلاسیک فرض بر این است که دقیق‌ترین فاصله اطمینان حاصل شود، یعنی فضای نمونه هم باید در نظر گرفته شود. اما در روش شواهدی از اصل درستنمایی تبعیت می‌کنیم که فقط به مشاهدات بستگی دارد. در پایان با استفاده از تکنیک شبیه سازی مونت کارلو این سه مقوله را با ارزیابی احتمال پوشش پارامتر واقعی مورد ارزیابی قرار می‌دهیم.

واژه‌های کلیدی: احتمال پیشین، احتمال پسین، تابع درستنمایی، احتمال پوشش.
کد موضوع‌بندی ریاضی (۲۰۱۰): 62C10، 62F25، 62F15.

---

۱ پیش‌گفتار

یکی از مهمترین تکنیک های آمار کاربردی استفاده از فاصله اطمینان است. متاسفانه در این روش به این نکته توجه نمی کنیم که امکان تکرار در آزمایشات وجود ندارد. زیرا از لحاظ نظری قضایای مربوطه متکی بر استنباط کلاسیک یا فراوانی گرا است. فاصله بیزی یا فاصله باور نیز به توزیع پیشین وابسته بوده که آن هم به اعتقادات شخصی تکیه می کند و نمی تواند نظر آمار دان یا محقق را جلب کند. براین اساس فاصله شواهدی که مبتنی بر تابع درستنمایی است را پیشنهاد می کنیم که هیچکدام از معایب روشهای فوق را ندارد.

توابع درستنمایی

فرض کنید متغیر تصادفي X داراي توزیع نرمال با میانگین μ و واریانس معلوم یك باشد. دو فرضیة رقیب زیر را در نظر مي‌گیریم.

براي بررسي دو فرضیة فوق به روش شواهدي نمونه‌اي تصادفي به حجم n اخذ مي‌کنیم. در اینصورت داریم

بررسي بهتر استنباط شواهدي را $Royall$ به طریق دیگري بیان مي‌کند که مي‌گوید "بجاي آنکه نسبت درستنمایي را براي دو فرضیة رقیب بدست آوریم، تابع درستنمایي را (باتقسیم بر ماکسیمم مقدار آن) طوري تعدیل مي‌کنیم که بیشترین مقدار آن یك باشد. سپس آن تابع را بر حسب پارامتر مورد

---

* نویسنده مسئول: emadi@um.ac.ir

$$\begin{cases} H_1 : \mu = 0 \\ H_2 : \mu = 1 \end{cases}$$

$$r = \frac{L_1}{L_2} = \frac{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-1}{2} x_i^2\right)}{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-1}{2} (x_i - 1)^2\right)}$$

$$= \prod_{i=1}^{n} \exp\left(\frac{1}{2} - x_i\right) = \exp\left(n\left(\frac{1}{2} - \bar{x}\right)\right),$$

نظر رسم کرده و در نهایت با توجه به منحني تابع درستنمايي استنباط شواهدي را نه تنها براي دو فرضية رقيب مورد بحث بلكه براي هر دو فرضية سادة رقيب انجام مي‌دهيم". اين روش براي مثال ۱-۲ به شكل زير است:

$$L(\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-1}{2} (x_i - \mu)^2\right)$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2} \sum_{i=1}^{n} (x_i - \mu)^2\right).$$

براي اينكه بيشترين مقدار $L(\mu)$ يك شود كافيست ماكسيمم $L(\mu)$ ر ا محاسبه كرده و $L(\mu)$ را بر آن تقسيم كنيم.

$$\max_{\mu \in R} L(\mu) = L(\bar{x})$$

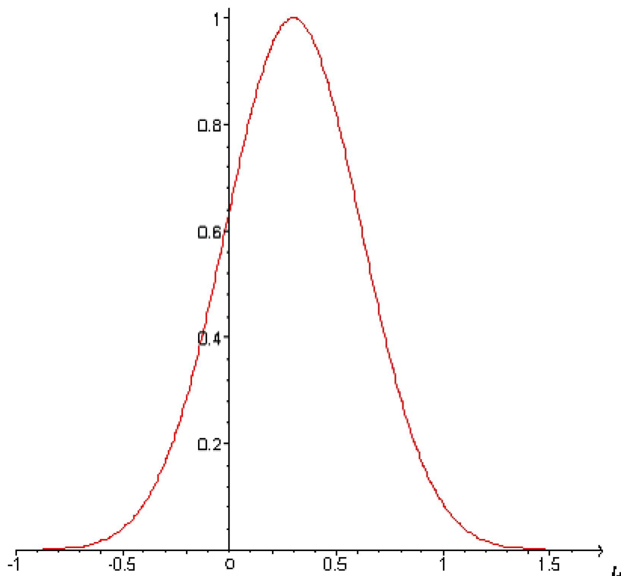
$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2} \sum_{i=1}^{n} (x_i - \bar{x})^2\right)$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2} \sum_{i=1}^{n} x_i^2 + \frac{n}{2} \bar{x}^2\right),$$

در نتیجه داریم

$$L(\mu) \propto \exp\left(-\frac{n(\mu-\bar{x})^2}{2}\right).$$

حال تابع فوق را بر حسب μ رسم مي‌کنيم.



شکل 1-1 نمودار تابع درستنمایی بر حسب μ بازای $n=10$، $\sigma^2=1$ (معلوم) و $\bar{x}=0.33$

با توجه به شکل ۱-۱ مشاهده مي‌شود $L(۰)$ بزرگتر از $L(۱)$ بوده پس داده‌ها از $H_۱$ پشتيباني بيشتري مي‌کنند و اگر k را برابر ۸ اختيار کنيم پشتيباني از $H_۱$ قوي محسوب مي‌شود.

فواصل آماری

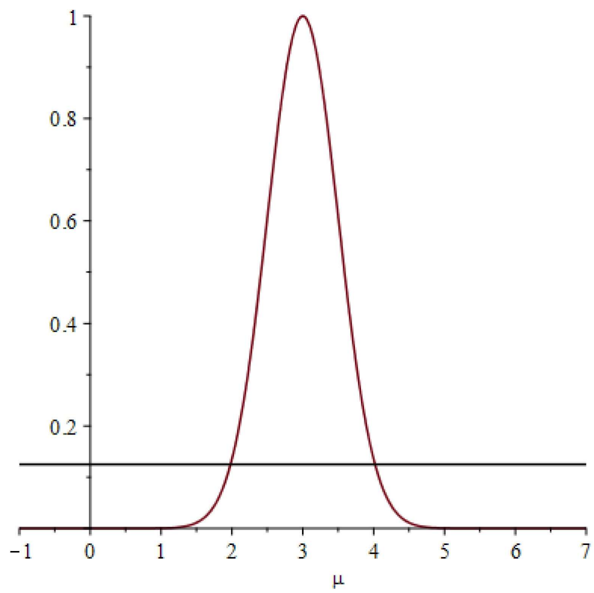
آمار کلاسیک: فاصله اطمینان

$$\bar{X} \pm \frac{\sigma}{\sqrt{n}} Z_{0.025}$$

آمار بیز: فاصله اعتبار یا باور

آمار شواهدی: فاصله شواهدی

$$\frac{n\bar{x}}{n+1} \pm \frac{Z_{0.025}}{\sqrt{n+1}}$$



تابع درستنمايي در قبال وجود پارامتر مزاحم

روش متعامدسازي

فرض کنید X_1 و X_2 متغيرهاي تصادفي مستقل از توزیع پواسن به ترتیب با پارامترهاي λ_1 و λ_2 باشند. تابع درستنمايي متعامد را براي پارامتر $\theta = \frac{\lambda_1}{\lambda_2}$ به دست مي‌آوریم براي این منظور پارامتر مزاحم λ را به صورت زیر تعریف مي‌کنیم

$$\gamma = \lambda_1 + \lambda_2.$$

با توجه به پارامتر مورد علاقة θ و پارامتر مزاحم λ داريم

$$\lambda_1 = \frac{\theta\gamma}{1+\theta}, \quad \lambda_2 = \frac{\gamma}{1+\theta},$$

تابع درستنمايي بر حسب پارامترهاي جديد عبارتست از

$$L(\theta,\gamma) \propto \frac{\theta^{x_1}}{(1+\theta)^{x_1+x_2}} \gamma^{x_1+x_2} e^{-\gamma}.$$

روش درستنمايي كناري

فرض كنيد $X_1, \cdots, X_n$ نمونه‌اي تصادفي از توزيع نرمال با پارامترهاي مجهول μ (ميانگين) و σ^2 (واريانس) باشند، و فرض كنيد كه پارامتر مورد علاقه باشد. پارامترهاي μ و σ^2 متعامد نيستند، زيرا تابع درستنمايي آنها به صورت زير است

$$L(\mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\right).$$

- روش درستنمايي نيمرخ

در اين روش ماكسيمم تابع درستنمايي را زماني كه مقدار پارامتر مورد توجه ثابت ($fixed$) در نظر گرفته شود، بدست مي‌آوريم. در اين حالت تابع درستنمايي مرتبط با پارامتر مورد علاقه را با نماد نمـايش مي‌دهيم (علامت $Profile$ است). مثال ۱-۵ فرض كنيد متغير تصادفي X داراي توزيع نرمال با ميانگين مجهول μ و واريانس σ^2 مجهول باشد. براي استنباط شواهدي دربارة پارامتر مورد علاقة μ بر اساس نمونه‌اي تصادفي به حجم n از توزيع فوق داريم.

$$L(\mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right).$$

روش درستنمايي شرطي

فرض كنيد متغيرهاي تصادفي مستقل X و Y داراي توزيع دوجمله‌اي به صورت زير باشند.

$$X \sim Bin(m, p_x), \quad Y \sim Bin(n, p_y),$$

كه m و n معلوم هستند. مي‌خواهيم استنباط شواهدي را براي نسبت بختها يعني

$$\psi = \frac{\frac{p_x}{1-p_x}}{\frac{p_y}{1-p_y}},$$

بررسي كنيم. براي اين منظور توزيع شرطي متغير تصادفي X به شرط متغير تصادفي $X+Y$ را بدست مي‌آوريم.

$$L_C(\psi) \propto \mathrm{P}\left(X = x \middle| X+Y = x+y;\ p_x, p_y\right)$$

$$\propto \frac{1}{\sum_{j=a}^{b} \binom{m}{j}\binom{n}{x+y-j}\psi^{j-x}},$$

روش درستنمايي برآورد شده

$$L(\mu,\sigma^2) \propto \sigma^{-n} \exp\left(-\frac{1}{2\sigma^2}\sum (x_i-\mu)^2\right)$$

$$L_E(\mu) \propto c(\underline{x}) s^{-n} \exp\left(-\frac{1}{2s^2}\sum (x_i-\mu)^2\right)$$

مراجع

عمادی، م. (۱۳۸۲)، پایان نامه دکتری، دانشگاه فردوسی مشهد.

Evidence interval as a substitute for confidence and credible intervals

Emadi, M.[1], **Mohammadi, M.**[2]
[1]Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran.
[2]Department of Statistics, University of Zabol, Zabol, Iran.

Abstract: One of the valid methods among statisticians to explain and interpret societies is to use interval estimation for the investigated parameters. Therefore, in this paper, while introducing the methods of obtaining interval estimation, we compare them by means of some examples. To introduce conventional methods, we can refer to the Bayesian method, the classical method, and the new evidence method. It is clear that each method has its own assumptions. For example, in the Bayesian method, it is assumed that the investigated parameter is random and has a probability function. In the classical method, it is assumed that the most accurate confidence interval is obtained, that is, the sample space should also be considered. But in the method of evidence, we follow the likelihood principle, which depends only on observations. In the end, using the Monte Carlo simulation technique, we evaluate these three methods by evaluating the probability of covering the real parameter.

Keywords: Prior probability, Posterior probability, Likelihood function, Coverage probability.
Mathematics Subject Classification (2010): 62C10, 62F25, 62F15.