اصول کلي نظريه و روشهاي استنباط آماري بيزي

دکتر محمد رضا مشکاني

گروه آمار ، دانشگاه شهيد بهشتي

مقدمه:

- پژوهش علمي (آزمايش ، مشاهده)

گردآوري داده ها : $D=\{x_1,...,x_n\}$

x_i نرده اي يا برداري

- روش آماري : آگاهي از سازوكار فرايند مولد D

- لزوم تدوين اصول موضوع براي هدايت فعاليت فوق

مسؤليت و وظيفه علم آمار در پژوهش هاي علمي
ياري رساندن به پژوهشگران تجربي

- تدوين پرسش اساسي پژوهش (فرمول بندي مسأله)
- شناسايي متغيرهاي مناسب (كنترل شده يا آزاد)
- شيوه درست گردآوري داده ها (طرح آزمايش ها يا نمونه گيري)
- خلاصه كردن داده ها (روش هاي توصيفي و تحليلي)
- استنباط آماري (تفسير پديده ها ، پيشگويي مقادير آتي)

مثال:

آتش سوزي در جنگلها : (عوامل مؤثر محيطي ، هواشناختي و مديريت)

احتمال رخ دادن آتش سوزي=p

$$p = f(h,t,x) = \frac{\exp(\alpha_1 h + \alpha_2 t + \alpha_3 x)}{1 + \exp(\alpha_1 h + \alpha_2 t + \alpha_3 x)}$$

<u>فايده اين مدل</u>: فروكاهي داده ها

تفسير آتش سوزي (در برابر تبيين آتش سوزي)

پيشگويي پيش آمدهاي آتي

لزوم مدل بندي احتمالاتي

عدم حتميت در داده ها و در نوع مدل پيشنهادي

مدل بندي پارامتري ، نيمه پارامتري ، ناپارامتري

مدل بندي پارامتري

$$D = (x_1, \dots, x_n), \quad x_i \sim \; p(x_i | \theta, x_{(-i)})$$

كه درحالت استقلال

$$(x_i | \theta) \overset{\text{iid}}{\sim} p(x_1 | \theta)$$

تعريف مدل آماري پارامتري:

X مشاهده مي شود ولي θ مشاهده شدني نيست و مجهول است

$$p(x|\theta), \;\; x \in \chi, \;\; \theta \in \Theta$$

وارونه سازي در استنباط آماري

$$\theta \to p(x|\theta) \to x$$

هدف از استنباط آماري آن است كه:

$$x \to \pi(\theta|x) \to \theta$$

ابزار لازم: قضيه بيز

$$\pi(\theta|x) \propto p(x|\theta) \cdot \pi(\theta), \quad \theta \in \Theta$$

تعريف مدل آماري بيزي:

مدل آماري $p(x|\theta)$ + توزيع پيشين $\pi(\theta)$

اصول فرو کاهي داده ها:

بايد داده هاي D را به فشرده ترين صورت ممکن فرو کاست.

چگونه ؟

با پيروي از اصول عقلاني و منطقي

اصل بسندگي (SP) ← آماره بسنده ملاك عمل است

اصل درستنمايي (LP) ← تابع درستنمايي راهنماست

اصل شرط مندي (CP) ← آنچه عملاً مشاهده شده مهم است

قضيه برون باؤم:

$$LP \Leftrightarrow SP + CP$$

با پيروي از اصل درستنمايي ، اصول ديگر نيز رعايت مي شوند

موضوع نظريه هاي بسامدي و بيزي در پيروي از اصول فروكاهي داده ها

- ✓ نظريه بسامدي: گاهي مي پذيرد ، گاهي عدول مي كند
- ✓ نظريه بيزي : همواره مي پذيرد

استنباط آماري و نظريه تصميم

ديدگاه ها :

برخي بسامد گرايان لزومي به پيروي از نظريه تصميم نمي بينند

برخي بسامد گرايان در قالب نظريه تصميم كار مي كنند

برخي بيزي گرايان لزومي به پيروي از نظريه تصميم نمي بينند

برخي بيزي گرايان در قالب نظريه تصميم كار مي كنند

- وضعيت غالب استفاده از نظريه تصميم در استنباط آماري است

هر گونه استنباطي يك تصميم است و مستلزم پيامدي است

كه ممكن است سود يا زيان داشته باشد.

ساختار نظریه تصمیم

- در هر مسأله بیش از یك تصمیم وجود دارد.
- تحت شرایط معیني هر دو تصمیم a_1 و a_2 قابل مقایسه اند

a_1 بهتر یا بدتر از a_2 است ، یا a_1 هم ارز a_2 است

اگر برتري را با نشان $\prec$ دهیم

$a_1 \prec a_2$ یا $a_2 \prec a_1$ یا $a_2 = a_1$

- اصل تراگذاري:

$$a_1 \prec a_2 \ \& \ a_2 \prec a_3 \Rightarrow a_1 \prec a_3$$

- اصل امر یقین:

$$[a_1 \prec a_2 | E, C] \& [a_1 \prec a_2 | E^C, C] \Rightarrow [a_1 \prec a_2 | C]$$

تابع مطلوبیت: به ازاي هر $\theta \in \Theta$، انتخاب تصمیم $a \in \mathrm{A}$ داراي پیامدي است که با تابع اندازه گیري مي شود. $u(c) = u(a,\theta)$

$u(\cdot,\cdot)$ تصمیم ها را به ازاي هر θ مرتب مي سازد.

در این کار باید عدم حتمیت مربوط به را نیز در نظر گرفت.

عدم حتمیت درباره θ :
$$\{\pi(\theta|c,a), \theta \in \Theta_a, a \in \mathrm{A}\}$$

از آنجا که متوسط مطلوبیت a عبارت است از

$$u(a|c) = \int_{\Theta_a} u(a,\theta)\pi(\theta|c,a)d\theta, \qquad a \in \mathrm{A}$$

که تصمیم ها را مرتب مي سازد.

بهتر و راحت تر است با تابع زیان نامنفي($L(a,\theta)$ کار کنیم.

$$L(a,\theta) = \sup_{a\in \mathrm{A}}\{u(a,\theta)\} - u(a,\theta)$$

(" تاوان" تصمیم نادرست به ازاي)= θ

متوسط زیان تصمیم ها را بر حسب زیان مرتب مي سازد.

$$L(a|c) = \int_{\Theta_a} L(a,\theta)\pi(\theta|c,a)d\theta, \qquad a \in \mathrm{A}$$

در مسأله هاي تصميم گيري ، رويكرد بالا مشكلي ندارد ، اما در استنباط آماري چطور؟

برخي معتقدند و استدلال مي كنند كه لزومي ندارد.

دو دليل عليه استدلال مخالفان:

(1) مسأله استنباط آماري وقتي ارزش تحليل داردكه درنهايت به تصميمي معقول بينجامد[رهزي: مشتي آرسنيك سمي است زيرا ممكن است كسي را بكشد ، نه به اين دليل كه عملاً كسي را كشته است]

(2) استنباط آماري ساختار رياضي مسأله تصميم را دارد.

فضاي كنش ها عبارت است از

$$A=\{\pi(\theta|D):\pi(\theta|D)\succ 0,\int_{\Theta}\pi(\theta|D)=1\}$$

$u(a,\theta)$ =(مطلوبيت a در برابر θ)،

توصيف عدم حتميت درباره θ = $\pi(\theta)$ و $a=\delta(D)$

$\theta\in\Theta$

پرسش هاي اساسي

اصولاً مفهومي به نام مطلوبيت وجود دارد ؟

(همين پرسش درباره زيان ناشي از يك تصميم مصداق دارد)

آيا توزيع پيشين $\pi(\theta)$ وجود خارجي دارد ؟چگونه تعيين مي شود ؟

پاسخ هر دو سؤال مثبت است.

پاسخ سؤال اول را كه از طريق اصل هاي موضوع مزبوط به پيامدهاي تصميم ها به دست مي آيد ، در De Groot(1970) فصل7 يا Robert(2001) فصل2 ببينيد.

پاسخ سؤال دوم ، از اصل تبادل پذيري مشاهدات به دست مي آيد/.

تعريف: مجموعه بردارهاي $\{x_1,\ldots,x_n\}$ تبادل پذير است اگر توزيع توأم آن ها تحت جايگشت هاي ممكن ناوردا باقي بماند.

دنباله $\{x_1,\ldots,x_n\}$ تبادل پذير است اگر هر زيرمجموعه متناهي از ان يك مجموعه تبادل پذير باشد.

قضيه نمايش دوفينتي

$$D=\{x_{i_1},\ldots,x_{i_n}\}\subset\{x_1,x_2,\ldots\}=Exchangebl\ e$$

آن گاه D " نمونه اي تصادفي" از مدل احتمالي

$$\{p(x|\omega),\omega\in\Omega\},x\in\chi$$

است كه با پارامتر ω توصيف مي شود و

$$\omega=\lim_{n\to\infty}f(D)$$

و اطلاعات موجود درباره مقدار ω در شرايط حاكم در C لزوماً با توزيع احتمال $\pi(\omega|C)$ توصيف مي شود.

اهميت قضيه نمايش دوفينتي

(1) اثبات وجود توزيع پيشيني براي هر مجموعه مشاهدات D

(2) اعطاي معنايي دقيق به پارامتر مجهول در مدل آماري

حالت ساده از قضيه نمايش

براي $\{x_i\}_{i=1}^{\infty} \sim Ber(w)$

يك توزيع احتمال F با تكيه گاه $[0, 1]$ وجود دارد به گونه اي كه

$$p_{n,k} = p\{x_1 = 1, \ldots, x_k = 1, x_{k+1} = 0, \ldots, x_n = 0\}$$

$$= \int_0^1 w^k (1-w)^{n-k} \, dF(w), \qquad \forall n, \quad 0 \le k \le n$$

برهان: دنباله متناهي $\{X_i\}_{i=1}^m$ را تبادل پذير بگيريد و قرار دهيد $q = p\{\sum_{j=1}^m x_j = r\}$

با اين شرط توزيع توأم $\{X_i\}_{i=1}^m$ مثل دو جمله اي منفي است.

$$p_{n,k} = \sum_{r=0}^{m} \frac{\binom{r}{k}\binom{m-r}{n-k}}{\binom{m}{n}\binom{n}{k}} q_r, \qquad 0 \le k \le n \le m$$

بگذاريد $m \to \infty$ و قرار دهيد $r = wm$ ، آن گاه

$$p_{n,k} = \int_0^1 \frac{\binom{wm}{k}\binom{(1-w)m}{n-k}}{\binom{m}{n}\binom{n}{k}} dF_m(w)$$

F_m تابع توزيعي با تكيه گاه $\{\frac{r}{m} : 0 \le r \le m\}$ و جهشهايي در $\frac{r}{m}$ با احتمال q_r است.

قضيه هاي اطمينان مي دهد كه $F_m \to F$ ، كه F توزيع يكنواخت خواهد داشت و

$$P_{n,k} \to \int_0^1 w^k (1-w)^{n-k} dF(w), \qquad \forall n, 0 \le k \le n$$

پيامد فرض تبادل پذيري

تبادل پذيري گسترش استقلال است

هر نمونه تصادفي (مستقل) نمونه اي تبادل پذير است

استقلال در نمونه شرطي و به شرط مقدار خاص پارامتر است و گرنه استنباط از $\{x_1, \ldots, x_n\}$

درباره متغير مستقل x_{n+1} مسخره است ، (ليندلي1972)

قضيه نمايش ، قضيه وجودي براي $\pi(w|c)$ روي Ω است و از هيچ ابزار ديگري

جز نظريه احتمال رياضي بهره نمي گيرد.

نتيجه: اگر دو مدل داراي پارامترهايي باشند كه به صورت تابعي با هم مرتبطند مقايسه توزيع هاي پسين آن ها معنايي عملي دارد.

مثال:

$$x \sim \mathrm{Pois}(\lambda), \quad \mathrm{E}(x \mid \lambda) = \lambda, \quad \lambda = \lim_{n \to \infty} \bar{x}_n$$

$$x \sim \mathrm{NB}(r, \theta), \quad \mathrm{E}(x \mid \theta, r) = \frac{r(1-\theta)}{\theta}, \quad \theta = \lim_{n \to \infty} \frac{r}{\bar{x}_n + r}$$

$$\Rightarrow \qquad \theta = \frac{r}{\lambda + r}$$

در توزيع پسين θ و $\frac{r}{\lambda + r}$ را بايد يكي انگاشت.

روش هاي استنباط بيزي

داده ها را با روش هاي توصيفي بررسي مي كنيم تا مدلي موقتي را پيشنهاد كنيم

فرض درست بودن اين مدل را با $\{\, p(D \mid w),\ w \in \Omega \,\}$

از دانش موجود D درباره وضعيت پارامتر $A\ k$ را در فضاي Ω پيش از به دست آوردن $\pi(w \mid k)$ مشخص مي كنيم.

از قضيه بيز توزيع پسين را به دست مي آوريم.

$$\pi(w \mid D) \equiv \pi(w \mid D, A, K) = \frac{p(D \mid w)\, \pi(w \mid k)}{\int_{\Omega} p(D \mid w)\, \pi(w \mid k)}, \qquad w \in \Omega$$

نکته: Ω فضايي است که $\pi(w|D) \succ 0$ است ، در غير اين صورت اگر $\pi(w|k) = 0$ به ازاي

، آن گاه $w \in \Omega_0$ به ازاي $\pi(w|D)=0$. پس همواره مي پذيريم که

$$\pi(w|k) \succ 0 \qquad w \in \Omega_0$$

هر پرسشي درباره w را به ياري $\pi(w|D)$ پاسخ مي دهيم

اين دستور کلي و جهان شمول استنباط بيزي است.

دیدگاه ها:

برخي معتقدند با به دست آوردن $\pi(w|D)$ كار ما پايان يافته است برخي برآنند كه

بايد بايد پاسخ پرسش هاي گوناگون را نيز فراهم كنيم

از جمله پرسش هاي مربوط به برآورد ، آزمون فرض ، پيشگويي ، رفتار مجانبي ، مدل

گزيني ،....

پرسشي مهم

آيا روش استنباط بيزي شيوه اي مشابه در ساير علوم دارد ؟

پاسخ: همة فعاليت هاي بشر در اندوختن علم و تجربه مبتني بر فرايند يادگيري است . فرايند يادگيري در روانشناسي به تفصيل تحت عنوان ادراك ، احساس ، شرطي شدن و غيره بحث مي شود. لب مطلب اين است.

پيش از مشاهده داده ها درباره يك موضوع ، فرد نسبت به آن موضوع به فراخور آگاهي ها و تجربيات ديگران و خودش فكر و انديشه و احساسي دارد[كه همان $\pi(w|k)$ است]پس از تجربة عملي و تماس با موضوع و گردآوري و مشاهدة داده ها ، در فكر خود بازنگري مي كند ، يا در انديشه و احساس خود راسختر مي شود ، يا آنرا تغيير مي دهد[كه همان $\pi(w|k)$ است] و روزآمد شدة $\pi(w|D,k)$ است

پيشگويي:

$$x_i|D \overset{\text{iid}}{\sim} p(x_i|w),\ i=1,\ldots,\ n$$

$$p(x_{n+1}|D) = ?\qquad D=\{x_1,\ldots,x_n\}$$

$$P(X_{n+1}|D)=\frac{p(x_{n+1},D)}{P(D)}=\frac{\int_\Omega p(x_{n+1},D|w)\pi(w)dw}{p(D)}$$

$$=\frac{\int_\Omega p(x_{n+1}|w)[p(D|w)\pi(w)dw]}{p(D)}$$

$$=\int_\Omega p(x_{n+1}|w)\pi(w|D)dw$$

مثال : كارخانه اي در $n=10$ ماه گذشته شكايتي راجع به كالايش دريافت نكرده است احتمال دريافت r شكايت در ماه آينده چقدر است ؟

$$x_i|\lambda \overset{iid}{\sim} Pois(\lambda), \quad \lambda \succ 0 \Rightarrow \quad p(D|\lambda) \propto \lambda^{n\bar{x}} e^{-n\lambda}$$

$$\pi(\lambda) = \lambda^{-1/2}$$

بنا به نبود اطلاعات:

$$\Rightarrow \pi(\lambda|D) \propto \lambda^{n\bar{x}-1/2} e^{-n\lambda} = Gam(n\bar{x}+1/2, n)$$

$$p(x_{n+1}|D) \propto \int_0^\infty (\lambda^{x_{n+1}} e^{-\lambda})(\lambda^{n\bar{x}-1/2} e^{-n\lambda}) d\lambda$$

$$= C \int_0^\infty \lambda^{x_{n+1}+n\bar{x}-1/2} e^{-(n+1)\lambda} d\lambda$$

$$= \frac{(n^{n\bar{x}+1/2})\Gamma(x_{n+1}+n\bar{x}+\frac{1}{2})}{x_{n+1}\Gamma(n\bar{x}+\frac{1}{2})(n+1)^{x_{n+1}+n\bar{x}+\frac{1}{2}}}, x_{n+1} = 0,1,\ldots$$

$$p(x_{n+1}=0|D) = 0.953, \quad p(x_{n+1}=1|D) = 0.043, \quad p(x_{n+1}=2|D) = 0.003$$

رفتار مجانبي:

پرسشي جالب: براي $D_n = \{x_1, \ldots, x_n\}$ و $\pi(w)$

اگر $n \to \infty$ $\quad \pi(w|D_n) \to ?$

اهميت پرسش: (1) تقريب مرتبة اوُل براي نتايج دقيق به ازاي n بزرگ

(2) روش هاي عيني بيزي از خواص مجانبي به دست مي آيند

دو حالت متمايز: (1) حالت گسسته:

$$\Omega = \{w_1, w_2, \ldots\}$$

w_t = مقدار واقعي و به فرض تمايز

$$p(x|w_i) \neq p(x|w_t), i \neq t$$

تحت شرايط نظم:

$$\lim_{n \to \infty} \pi(w_i|D_n) = \begin{cases} 1 & i = t \\ 0 & i \neq t \end{cases}$$

(2)حالت پیوسته: $\Omega \subset R^K$ ، با بسط $\pi(w|D_n)$ پیرامون $\hat{w} = MLE(w)$ و به فرض شرایط نظم
(به كمك قضية حد مركزي و قانون اعداد بزرگ)

$$\pi(w|D_n) \xrightarrow{D} N_k(\hat{w}, \tfrac{1}{n} F^{-1}(\hat{w}))$$

$F(w)$ ماتریس اطلاع فیشر است.

مثال: $x_i \sim Ber(w)$

$$\pi(w|n,r) \xrightarrow{D} N(\tfrac{r}{n}, \tfrac{1}{n}\hat{\theta}(1-\hat{\theta}))$$

برآورد نقطه اي:

براي هر تابع پارامتري $\theta = \theta(w) \in \Theta$ و در دست داشتن $\pi(w|D)$

با فرض $L(\tilde{\theta},\theta)$

$$\theta^* = \arg\inf_{\tilde{\theta}} \int_{\Theta} L(\tilde{\theta},\theta)\,\pi(\theta \mid D)\,d\theta$$

$$= \arg\inf_{\tilde{\theta}} \int_{\Omega} L(\tilde{\theta}(w),\theta(w))\,\pi(w \mid D)\,|J|\,dw$$

زيانهاي متداول و برآوردهاي متناظر

$$L(\tilde{\theta},\theta) = (\tilde{\theta}-\theta)'(\tilde{\theta}-\theta) \qquad \Rightarrow \qquad \theta^* = E(\theta|D)$$

به ازاي B_ε، كره اي به مركز θ و شعاع ε و :

$$L(\tilde{\theta},\theta)=\begin{cases}1, & \tilde{\theta}\in B_{\varepsilon}\\ 0, & o.w.\end{cases} \quad \Rightarrow \quad \theta_{\varepsilon}^{*} \rightarrow Mode\ of\ (\theta\,|\,D) \quad as\ \varepsilon\rightarrow 0$$

$$L(\tilde{\theta},\theta)=\begin{cases}c_1(\tilde{\theta}-\theta) & \tilde{\theta}\geq\theta\\ c_2(\tilde{\theta}-\theta) & o.w.\end{cases}$$

[$\Rightarrow \theta^{*}=$ چندك $\frac{c_2}{c_2+c_1}$ ام از توزيع $\pi(\theta|D)$]

$$c_2=c_1 \quad \Rightarrow \quad \theta^{*}=Median\ of\ (\theta\,|\,D)$$

در اين حالت ها ناوردايي در تبديل پارامتر برقرار نيست:

$$\phi^{*}(\theta)\neq\phi(\theta^{*})$$

مثال:

براي $x \sim Bin(n,\theta)$ ، $D=(n,r)$ ، $\theta \sim Bet\,(\frac{1}{2},\frac{1}{2})$ $L(\tilde{\theta},\theta)=(\tilde{\theta}-\theta)^2$

خواهيم داشت:

$$\theta_q^{\ *} = \frac{r+\frac{1}{2}}{n+1}$$

اگر تبديل $\phi(\theta)=\log\{\frac{\theta}{1-\theta}\}$ را به كار ببريم

$$\phi_q^{\ *}(\theta)=\psi(r+\frac{1}{2})-\psi(n-r+\frac{1}{2}), \qquad \psi(\Im)=\frac{d}{d\Im}\log(\Gamma(\Im))$$

$$\phi_q^{\ *}(\theta)\neq\phi(\theta_q^{\ *})$$

برآوردگر ناوردا :

اختلاف ذاتي دو توزيع $p_2(x)$ و $p_1(x)$

$$\delta\{p_1, p_2\} = \min\left\{\int_{X} p_1(x)\log\left(\frac{p_1(x)}{p_2(x)}\right)dx, \int_{X} p_2(x)\log\left(\frac{P_2(x)}{p_1(x)}\right)dx\right\}$$

متقارن ، نامنفي ، صفر اگر وتنها اگر $p_1(x) = p_2(x)$ به ازاي $x \in X$

تابع زيان ذاتي:

$$\delta(\tilde{\theta}, \theta) = \min\{\Re(\tilde{\theta} \mid \theta), k(\theta \mid \tilde{\theta})\}$$

$$\Re(\theta_i, \theta_j) = \int_{\tau} p(t \mid \theta_j)\log\frac{p(t \mid \theta_j)}{p(t \mid \theta_i)}dt$$

هر آماره بسنده حتي خود D

$$= t = t(D) \in \tau$$

با وجود پارامتر مزاحم : λ

$$\delta\{\tilde{\theta},(\theta,\lambda)\} = \min_{\lambda_i \in \Lambda} \delta\{(\tilde{\theta},\lambda_i),(\theta,\lambda_i)\}$$

امید رياضي پسين اختلاف ذاتي دو توزيع $p(x \mid \tilde{\theta})$ و $p(x \mid \theta)$

$$d(\tilde{\theta} \mid D) = \int_\Lambda \int_\Theta \delta\{\tilde{\theta},(\theta,\lambda)\}\, \pi(\theta,\lambda \mid D)\, d\theta\, d\lambda$$

آن مقدار θ^* که $d(\tilde{\theta}, D)$ را مينيمم سازد برآوردگر ذاتي است.

برآوردگر ذاتي ناورداي بيزي:

$$\theta^* = \theta^*(D) = \arg \min_{\tilde{\theta} \in \Theta} d(\tilde{\theta}, D)$$

مثال:

$$\delta(\theta_1,\theta_2 \mid n) = n \min\{\Re(\theta_1 \mid \theta_2), k(\theta_2 \mid \theta_1)\}$$

$$\Re(\theta_i \mid \theta_j) = \theta_j \log\left(\frac{\theta_i}{\theta_j}\right) + (1-\theta_j)\log\left(\frac{1-\theta_j}{1-\theta_i}\right)$$

$$\theta^* = \arg \min_{\tilde{\theta} \in \Theta} \; d(\tilde{\theta} \mid D)$$

$$= \arg \min_{\tilde{\theta} \in \Theta} \int_0^1 \delta(\tilde{\theta}, \theta \mid n)\, \pi(\theta \mid n, r)\, d\theta$$

به طور عددي بايد حل كرد.

يك تقريب خوب

$$\theta_I^* \approx \frac{r + \frac{1}{3}}{n + \frac{2}{3}}$$

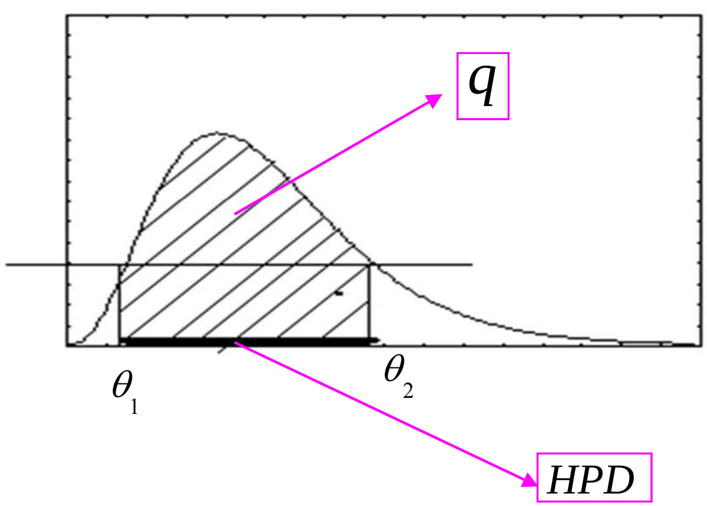
برآورد ناحيه اي(بازه اي): به ازاي $0 \prec q \prec 1$ پيدا كنيد $R_q \subset \Theta$

به طوري كه:

$$\int_{R_q} \pi(\theta \mid D)\, d\theta = q$$

$HPD =$ ناحيه اي مانند R_q كه در آن به ازاي هر $\theta_i \in R_q$ و $\theta_j \notin R_q$

$\pi(\theta_i \mid D) \succ \pi(\theta_j \mid D)$ و در ضمن شرط بالا برقرار است



مثال را در بخش كاربرد خواهيم ديد

آزمون فرض:

$$\begin{cases} H_0 \equiv \{\theta \in \Theta_0 \subset \Theta\} \\ H_1 \equiv \{\theta \notin \Theta_0\} \end{cases}$$

$a_0 =$ اعلام کنید H_0 درست است

$a_1 =$ اعلام کنید H_0 درست است $\quad \Longrightarrow \quad \mathrm{A} = \{a_0, a_1\}$

تابع زیان $L(a_i, \theta)$ را مشخص کنید

مثلاً تابع زیان $1-0$ چنین است

$$\begin{cases} \{L(a_0,\theta)=0\,,\, L(a_1,\theta)=1\} & ,\theta \in \Theta_0 \\ \{L(a_1,\theta)=1\,,\, L(a_1,\theta)=0\} & ,\theta \notin \Theta_0 \end{cases}$$

مزیت رد کردن H_0

$$\Delta L(\theta) = L(a_0,\theta) - L(a_1,\theta) = \begin{cases} -1 \;, & \theta \in \Theta_0 \\ +1 \;, & \theta \notin \Theta_0 \end{cases}$$

H_0 را رد کنید اگر

$$p(\theta \notin \Theta_0 \mid D) \succ p(\theta \in \Theta_0 \mid D) = 1 - P(\theta \notin \Theta_0 \mid D)$$

$$\Rightarrow \qquad P(\theta \notin \Theta_0 \mid D) \succ \frac{1}{2}$$

در حالت $\Theta_0 = \{\theta_0\}$ و θ پیوسته $P(\theta \in \Theta_0) = 0$ مشکل زاست

در این به θ_0 جرمي مانند P_0 اختصاص مي دهيم

در این جا هم از روش آزمون فرض ذاتي مي توان بهره گرفت.

مثال را در بخش کاربرد خواهيم ديد

تحليل مرجعي: $p(x|\theta)$ ، $\pi(\theta)$ مؤلفه هاي اساسي اند

انتخاب توزيع $\pi(\theta)$ چگونه صورت مي گيرد ؟

به كوتاهي

$\pi(\theta)$ دو جنبه دارد: فرم رياضي تابع $\pi(.)$ و گاهي در صورت وجود پارامترهاي آن بايد تعيين شوند. ذهنيت پژوهشگر در هر دو جنبه دخيل است.

(1) **توزيع هاي پيشين "عيني"**: آن توزيع هايي كه داده ها ديكته مي كنند:

پيشين جفريز ، پيشين مرجع

(2) **توزيع هاي پيشين "نيمه عيني"**: آن توزيع هايي كه فرم رياضي آنها از $p(x|\theta)$ الهام گرفته مي شود و پارامترهاي آنرا پژوهشگر انتخاب مي كند.

توزيع هاي پيشين مزدوج ، پيشين هاي مزدوج برآورد شده(بيز تجربي)

(3) **توزيع هاي پيشين ذهني**: هم فرم $\pi(\theta)$ و هم پارامترهاي آن را پژوهشگر تعيين مي كند.

از كدام دسته در چه موقعيتي استفاده كنيم ؟ بستگي به ميزان اطلاعات پيشين دربارة θ دارد.

روش پيشين عيني را تحليل مر جعي گويند.

الگوريتم هاي تعيين پيشين مرجع: آن پيشيني را كه اطلاعات گمشده را ماكسيمم مي سازد ، انتخاب كنيد.

آمارة بسنده: $t = t(D) \in \tau$ ، $\theta \in \Theta$ ، $p(D \mid \theta)$

ميزان اطلاع بر اساس نظريه

اطلاع شانون

$$= I^{\theta}\{\tau, p(\theta)\} = \int_{\tau}\int_{\Theta} p(t,\theta)\log\left(\frac{p(t,\theta)}{p(t)\ p(\theta)}\right) d\theta\ dt$$

$$= E_t\Big[\int_{\Theta} p(\theta \mid t)\log\left(\frac{p(\theta \mid t)}{p(\theta)}\right) d\theta\Big]$$

$$= [\ p(\theta) \qquad\qquad]$$

تابعي از

ابتدا به ازاي $(t_1, \ldots, t_k)$ ، $\pi(\theta)$ را حساب كنيد و بعد $I^{\theta}\{\tau^K, P(\theta)\}$ جواب حاصل پيشين مرجع $k \to \infty$ است.

در حالت برداري θ ، اين عمليات گام به گام براي هر θ_j انجام مي شود.

<u>ارتباط بين شيوه هاي بهينه بسامدي و شيوه هاي بيزي:</u>

قضايايي چند رابطة بين تصميمهاي مينيماكس و پذيرفتني را برقرار مي سازند

نتيجه: تحقق ملاكهاي بهينگي تصميم گيري در آمار بسامدي را از طريق برآوردگردهاي بيزي قابل دسترسي است.

مثال: آيا T^2 ي هتلينگ ، $\mathrm{T}^2 = n\,(\bar{x}-\mu)'\,S^{-1}\,(\bar{x}-\mu)$ آزموني پذيرفتني براي $H_0 : \mu = \mu_0$ در توزيع

$$N_p(\mu,\Sigma) \qquad \theta = (\mu,\Sigma)$$

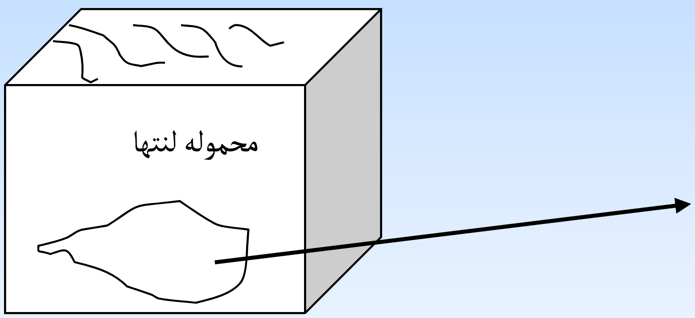
است ؟ از راه يافتن پيشيني كه به اين ملاك بيانجامد ، مي توان نشان داد كه پذيرفتني است.

مثال كاربردي

شركت A: لنت ترمز مي سازد.

شركت B: خودرو ساز است و لنتهاي A را براي خودروهايش مي خرد.

مي خواهند با هم قرارداد ببندند. طول عمر لنتها و در صد لنتهاي زير استانداردار نظر حقوقي و بازرگاني براي دو طرف مهم است.



نمونة تصادفي $= D = \{x_1, \ldots, x_n\}$

آزمايش مخرب است و تا از بين رفتن نمونه ادامه مي يابد $x =$ طول عمر لنت

$$(x_i \mid \theta) \overset{\text{i.i.d.}}{\sim} p(x_i \mid D) = \theta e^{-\theta x_i} \qquad , x_i \succ 0 \,, \theta \succ 0$$

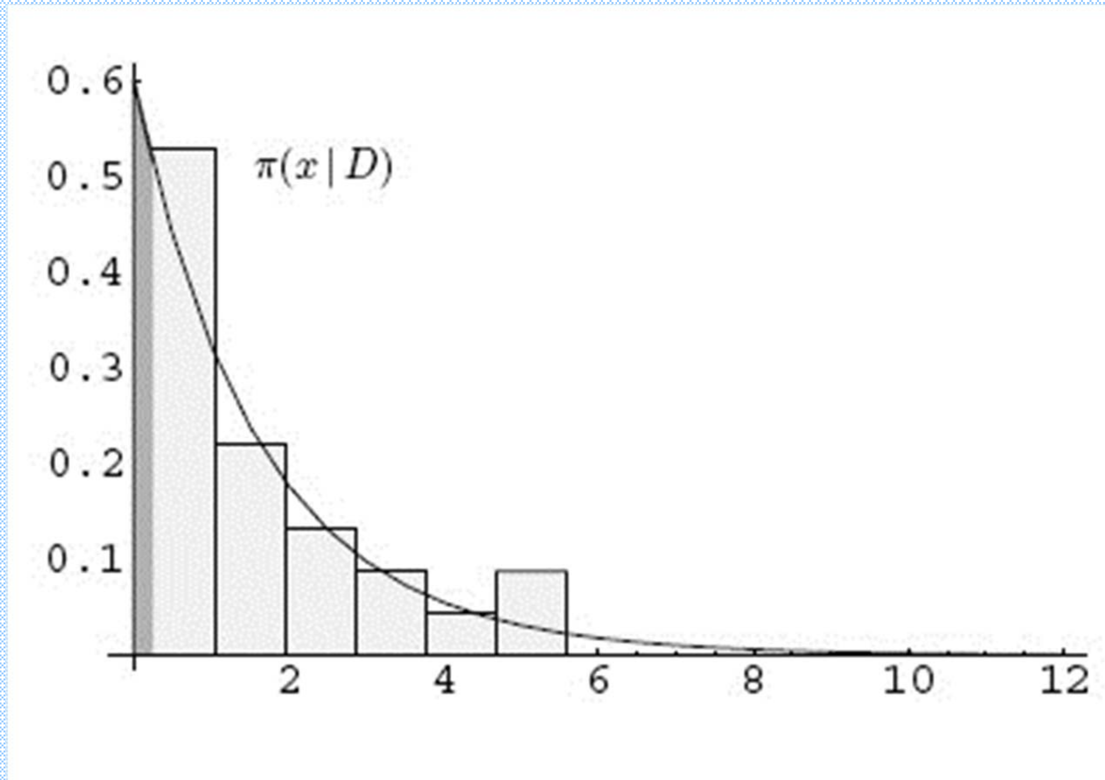
موارد مورد نظر:

نرخ خطر $= \theta = \dfrac{p(x \mid \theta)}{1 - F(X \mid \theta)}$

طول عمر لنت هاي بعدي $= x_{n+1}$

$n = 25$ شكل 1 ، توزيع شرطي نمايي را براي

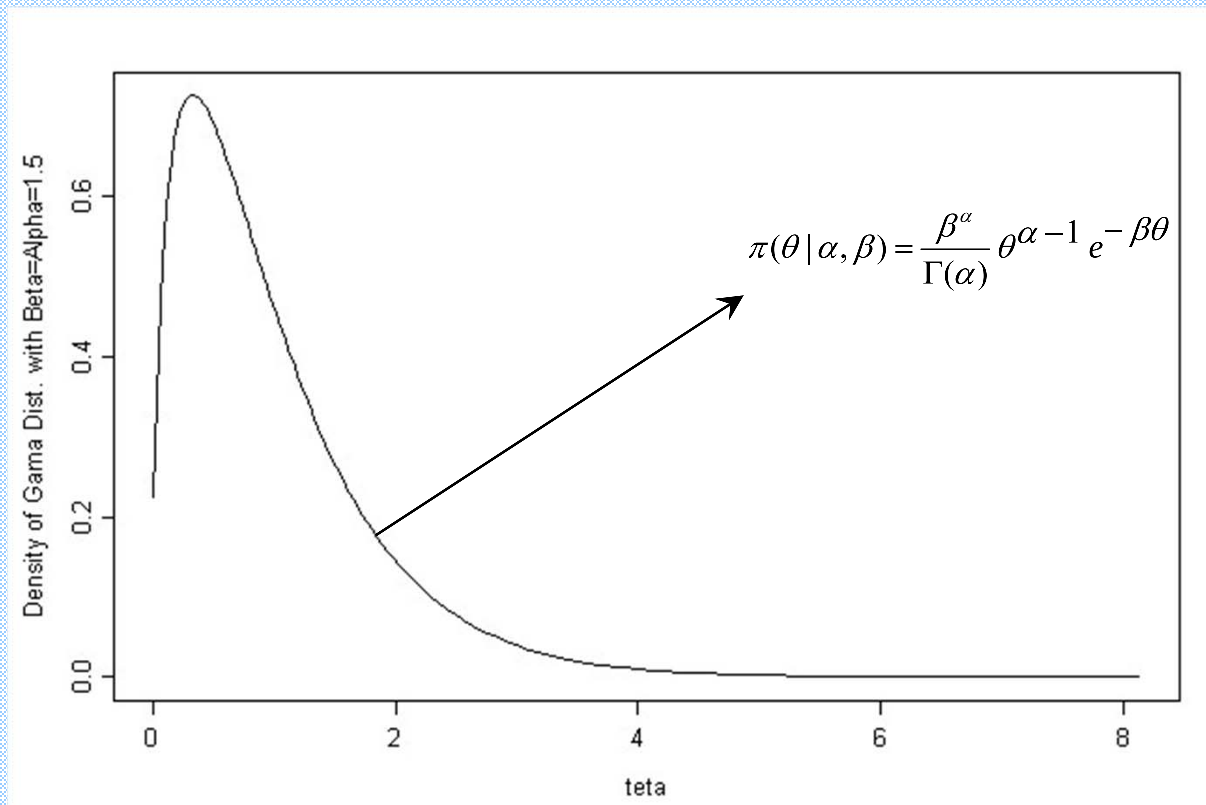
لنت پيشنهاد مي كند



شكل1. بافت نگار داده ها و مدل نمايي پيشنهادي براي $p(x \mid D)$

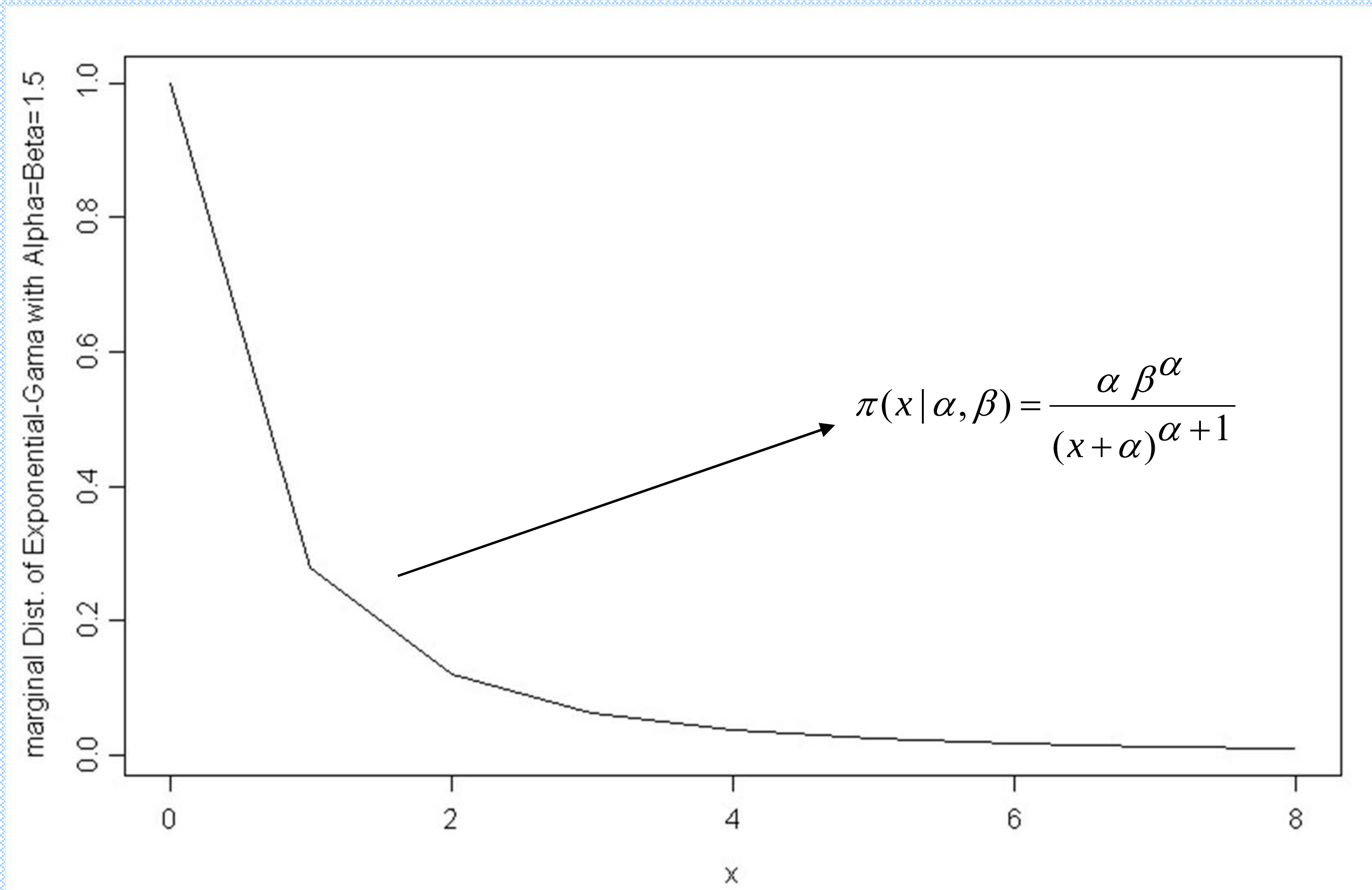
براي θ پيشين $\alpha, \beta \succ 0$ $Gam(\alpha,\beta)$ پيشنهاد

مي شود{ $\pi(\theta \mid \alpha,\beta) = Gam(\alpha,\beta)$ }



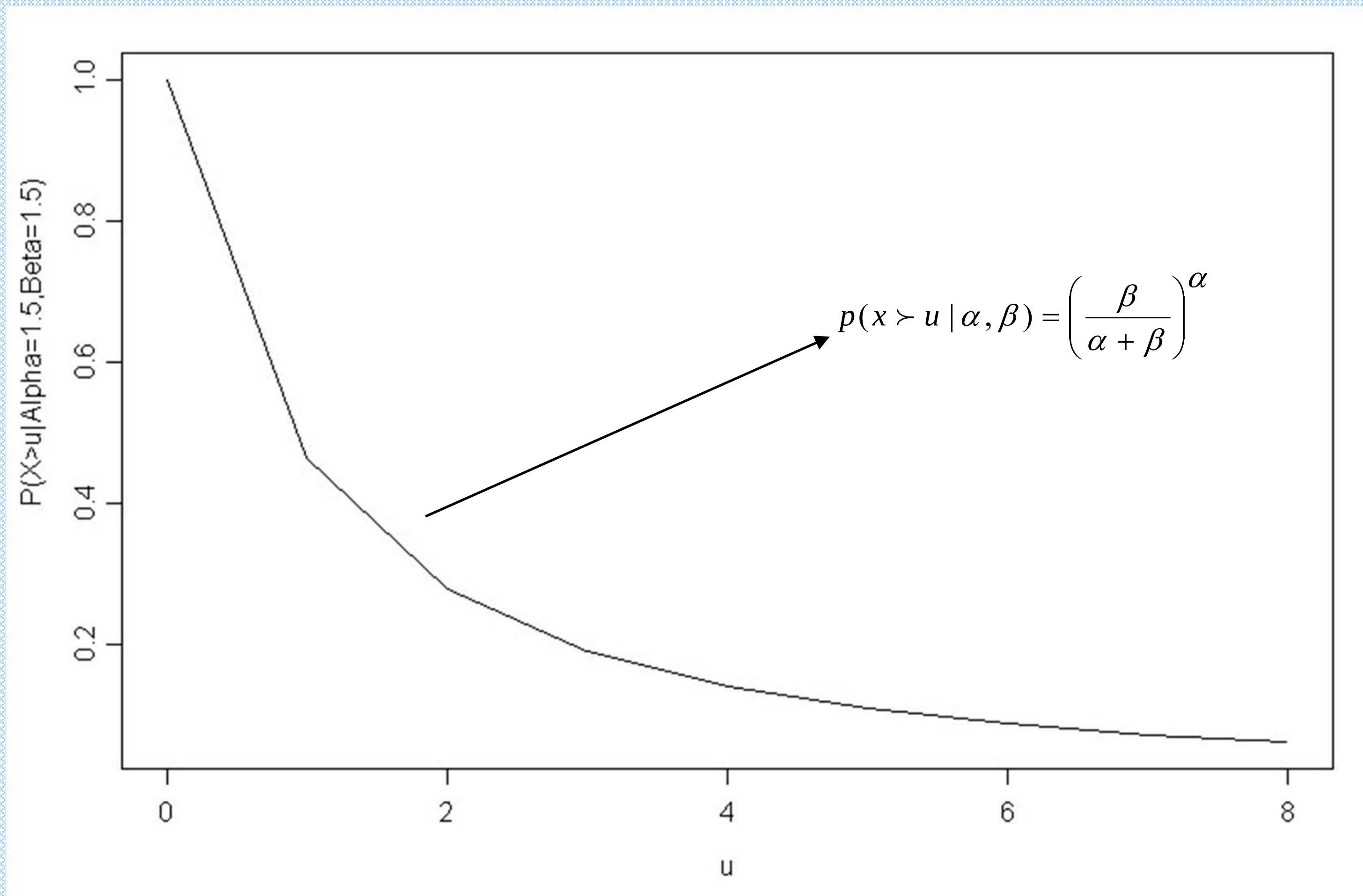
شكل2. توزيع پيشين گاما براي θ : $\pi(\theta \mid \alpha,\beta) = Gam(\alpha,\beta)$

شکل3 ، توزيع حاشيه اي نمايي-گاما براي X را نشان مي دهد



شکل3. توزيع حاشيه اي نمايي-گاما براي X

شکل4 ، نشانگر تابع قابلیت اعتماد است



شکل4. تابع قابلیت اعتماد

داده ها بر حسب هزار ساعت:

$$s=\sum_{j=1}^{n} x_j = 41.574$$

$$\bar{x} = s/25 = 1.663$$

$$\hat{\theta} = MLE = \frac{1}{1.663} = 0.601$$

$$SD(x) = 1.286 \qquad , \quad Range = [0.136\,,\,5.591]$$

خصوصيات شبيه به توزيع نمايي است.

$$\pi(\theta \mid D) \approx N(\hat{\theta}\,,\frac{1}{n}\hat{\theta}^2) = N(0.601\,,0.120^2)$$

بازة اطمينان:

$$0.601 \pm 1.96\,(0.120) \quad \Rightarrow \quad (0.366\,,\,0.837)$$

تعيين پيشين مرجع: $F(\theta)=\theta^{-2}$ پس $\pi(\theta) \propto \theta^{-1}$

$$\pi(\theta \mid D) \propto \theta^{n-1} e^{-s\theta} \sim \mathrm{Gam(n,\ s)} = (25,\ 41.57)$$

$$\mathrm{E}(\theta \mid \mathrm{D}) = \mathrm{MLE}(\theta) = 0.601$$

$\theta_{\mathrm{n}} =$ مد پسيني $= \dfrac{n-1}{s} = 0.577$

$SD(\theta) =$ انحراف معيار پسيني $= \dfrac{\sqrt{n}}{s} = 0.120$

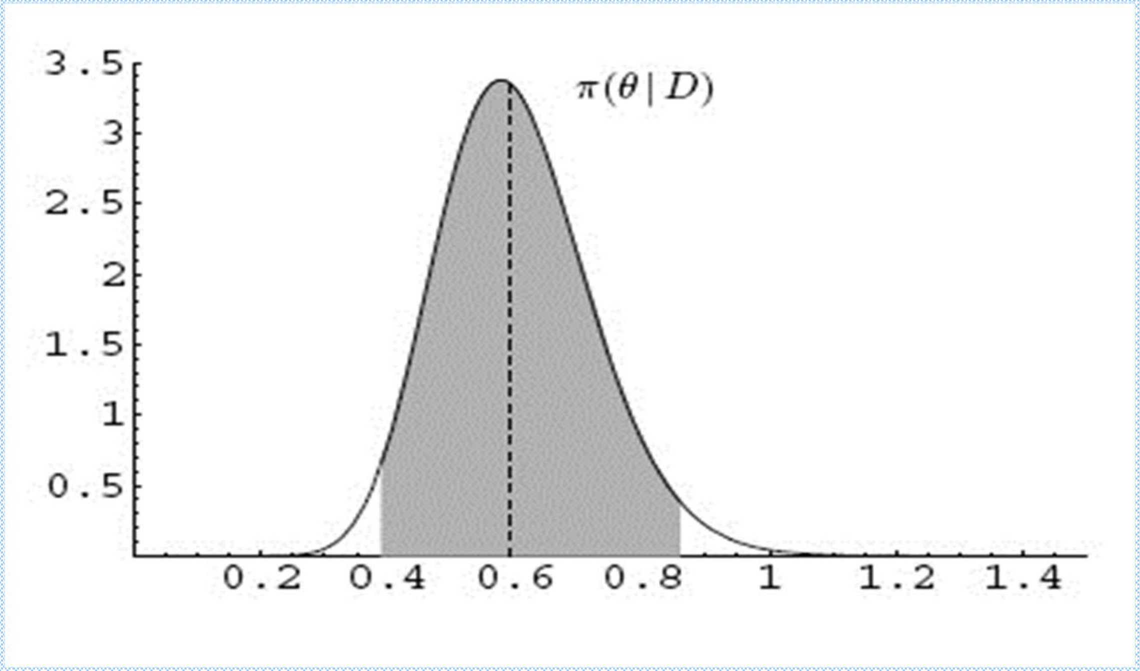
انتگرال گاماي ناقص $P = P(\theta \prec a \mid D) = \int_0^a Gam(n,s)\, d\theta =$

α	0.389	0.593	0.859
β	0.025	0.500	0.975

بازة باوركردني 95 در صدي بسيني: $(0.389 \quad , 0.859 \quad)$

ميانه پسيني $= 0.593$

شكل 5. توزيع پسيني مرجع نرخ خطر θ . ناحية هاشور خورده بازة باور كردني 95در صدي است. خط نقطه چين مكان برآوردگر ذاتي را نشان مي دهد.



شكل 5. توزيع پسيني مرجع نرخ خطر θ

توزيع پيشگو

$$p(x \mid D) = \int_0^{\infty} p(x \mid D)\; p(\theta \mid D)\, d\theta$$

$$= \frac{n\, s^n}{(x+s)^{n+1}}$$

$$p(x \prec b \mid D) = \int_0^{b} \frac{n\, s^n}{(x+s)^{n+1}}\, dx \;=\; 1 - \left(\frac{s}{s+b}\right)^n$$

استاندارد قابل قبول=250 ساعت

در آينده 14% لنت ها زير استاندارد خواهد بود

$$p(x \prec 0.250 \mid D) = 0.139 \;\approx\; \%14$$

برآورد نقطه اي در تنظيم قرارداد ضروري است.

$$\theta^* = \arg \min_{\tilde{\theta}} \left[d(\tilde{\theta} \mid n,s) = n \int_0^{\infty} \delta(\tilde{\theta},\theta)\, Gam_{\theta}(n,s)\, d\theta \right]$$

محاسبات عددي لازم است.

$$\theta^* = 0.5899$$

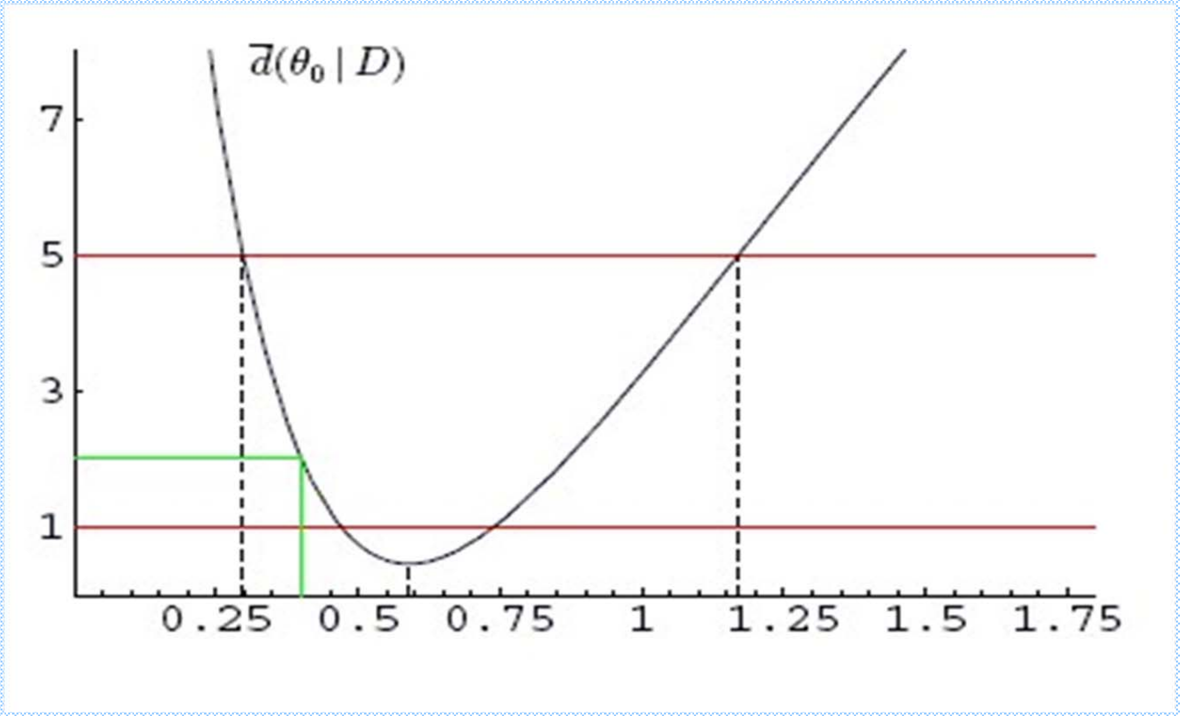
آزمون فرض براي آگهي هاي تبليغاتي لازم است.نبايد مقداري ازθ ذكر شود كه باعث شكايت و وارد آمدن خسارت به شركتA شود.

$$H_0 : \theta \prec 0.4$$

يعني متوسط طول عمر لنتها $2.5 = \frac{1}{0.4}$ (هزار ساعت)=2500ساعت است.

$$p\left[\theta \prec 0.4 \mid D\right] = 0.033$$

شکل6. امید رياضي زيان ذاتي پسين برايθ_0 به عنوان جانشيني براي θ ي واقعي مينيمم در نقطة $\theta^* = 0.590$ است. مقادير خارج از بازة $(0.297\ ,1.171)$بنابر قرارداد مقادير ردکردني اند

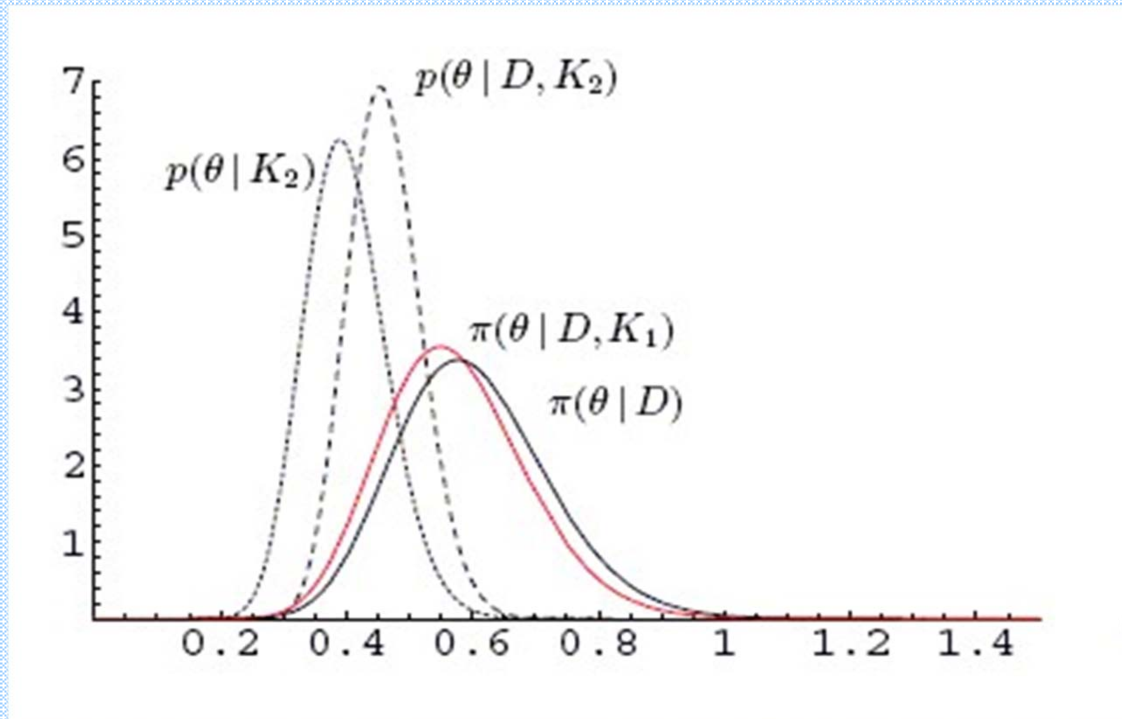


شکل6. امید رياضي زيان ذاتي پسين براي θبه عنوان جانشيني برايθ_0واقعي

شکل7. چگاليهاي احتمال نرخ خطر θ ، تحت مفروضات مختلف براي

... = پيشين ذهني ، --- = پسين ذهني ،

—— = پسين مرجع جزءاً اگاهي بخش ، = پسين مرجع قراردادي



شکل7. چگاليهاي احتمال نرخ خطر θ

نتيجه

- استنباط بيزي توانايي فوق العاده براي حل مسأله هاي مختلف آماري دارد
- مفاهيم آن بسيار ساده و عقلاني اند
- احتياج به مفاهيم اضافي غير از مفاهيم نظريه احتمال نيست
- الگوريتم آن بسيار ساده است ولي محاسبات مي تواند مشكل باشد

$P(x|\theta)$ را مشخص كنيد

$\pi(\theta)$ را انتخاب كنيد

$\pi(\theta|x) \propto p(x|\theta)\,\pi(\theta)$ را به دست آوريد

از اين توزيع هر سؤالي را راجع به θ پاسخ دهيد.

اگر تابع زيان $L(\delta,\theta)$ را به كار مي بريد

را حساب كنيد.

$$\tilde{\theta} = \arg\min_{\delta} E[L(\delta,\theta)|D]$$

با تشکر