برآورد گر های همپایا

Equivaricance Estimators

Http://MiladZaman005.blogfa.com

نویسنده : سید جمال میرکمالی

منبع: **E.L.Lehmann,G.Casella, Theory of Point Estimation, Second Edition,Springer-Verlag,1998**

گروه دانش آماری - بهمن ۱۳۸۵

اصول پایایی برآورد گر ها

در حالت کلی یک نمونه تصادفی $\mathbf{X}=(X_1,...,X_n)$ با تابع چگالی توام $f_{\mathbf{X}}(\mathbf{x})$ داریم. بوسیله یک تبدیل $\mathbf{Y}=g(\mathbf{X})$ داده های جدید را تولید می کنیم.این تبدیل روی خانواده توزیع های $\mathcal{P}=\{f_\theta:\theta\in\Theta\}$ یک گروه از تبدیلات بوجود می آورد. این گروه از تبدیلات را با $\mathcal{G}$ نمایش می دهیم.

پایایی خانواده توزیع ها: فرض کنید $\mathcal{P}=\{f_\theta:\theta\in\Theta\}$ یک خانواده از توزیع های احتمال برای X باشد و $\mathcal{G}$ یک گروه از تبدیلات روی $\boldsymbol{X}$ باشد. خانواده $\mathcal{P}$ را تحت گروه $\mathcal{G}$ پایا گوییم اگر برای هر $\theta\in\Theta$ ، یک $\theta'\in\Theta$ یکتا وجود داشته باشد بطوریکه $\mathbf{Y}=g(\mathbf{X})$ دارای تابع چگالی $f_{\theta'}$ باشد . در این صورت θ' را با $\bar{g}(\theta)$ نمایش می دهیم و تبدیلات بوجود آمده روی فضای پارامتر تحت گروه تبدیلات $\mathcal{G}$ را $\bar{\mathcal{G}}$ می نامیم.

پایایی تابع زیان: فرض کنید خانواده $\mathcal{P}=\{f_\theta:\theta\in\Theta\}$ تحت گروه تبدیلات $\mathcal{G}$ پایا باشد. در این صورت تابع زیان $L(\theta,\delta)$ را تحت گروه $\mathcal{G}$ پایا گوییم اگر برای هر $g\in\mathcal{G}$ و هر $\delta\in\mathcal{D}$ (فضای برآوردگر ها $=\mathcal{D}$) یک $\delta^*\in\mathcal{D}$ یکتا وجود داشته باشد بطوریکه

$$L(\theta,\delta)=L(\bar{g}(\theta),\delta^*)$$

در این حالت δ^* را با $\tilde{g}(\delta)$ نمایش می دهیم و تبدیلات بوجود آمده روی فضای برآورد گرها تحت گروه تبدیلات $\mathcal{G}$ را $\tilde{\mathcal{G}}$ می نامیم.
اگر در مساله استباط آماری ، هم خانواده توزیع ها و هم تابع زیان پایا باشد ، گوییم مساله تحت گروه تبدیلات $\mathcal{G}$ پایا است.

اصول پایایی :

۱- پایایی اندازه گیری (تابعی): اگر $\delta(\mathbf{X})$ را برای برآورد θ بکار بریم ، طبیعی است $\tilde{g}(\delta(X))$ را برای برآورد $\bar{g}(\theta)$ بکار بریم.

۲- پایای ساختاری (رسمی): اگر دو مساله استنباط آماری یک ساختار از لحاظ پارامتر ، تابع چگالی و تابع زیان داشته باشند ، طبیعی است روش یکسانی را برای دو مساله بکار بریم ؛ یعنی $\delta(g(\mathbf{X}))$ را برای برآورد $\bar{g}(\theta)$ بکار بریم.

برآورد گر همپایا: فرض کنید مساله مورد نظر ما تحت گروه تبدیلات $\mathcal{G}$ پایا باشد. برآوردگر $\delta(\mathbf{X})$ را تحت گروه تبدیلات $\mathcal{G}$ همپایا گوییم هرگاه برای هر $g\in\mathcal{G}$ داشته باشیم $\delta(g(\mathbf{X}))=\tilde{g}(\delta(\mathbf{X}))$.

بهترین برآوردگر همپایا: برآورد گر همپایایی که دارای کمترین مخاطره باشد ، بهترین برآورد گر همپایا MRE نامیده می شود.

لم : اگر خانواده $\mathcal{P}$ تحت گروه تبدیلات $\mathcal{G}$ پایا باشد آنگاه

$$E_\theta[h(g(\mathbf{X}))]=E_{\bar{g}(\theta)}[h(\mathbf{X})]$$

قضیه : فرض کنید δ یک برآورد گر همپایا در یک مساله پایا تحت گروه تبدیلات $\mathcal{G}$ باشد. همچنین فرض کنید برای هر دو نقطه $\theta,\theta'\in\Theta$ ، تبدیل $\bar{g}\in\bar{\mathcal{G}}$ وجود داشته باشد بطوری که $\theta'=\bar{g}(\theta)$ ؛ در این صورت $R(\theta,\delta)=R(\bar{g}(\theta),\delta)=R(\theta',\delta)$. یعنی مخاطره نسبت به θ ثابت است.

گروهی از تبدیلات را که برای هر دو نقطه $\theta,\theta'\in\Theta$ تبدیل $\bar{g}\in\bar{\mathcal{G}}$ وجود داشته باشد بطوری که $\theta'=\bar{g}(\theta)$؛ یک گروه انتقالی $transitive$ گویند.

MRE در خانواده مکانی توزیع ها:

فرض کنید $\mathbf{X}$ دارای چگالی $f_\theta(\mathbf{x})=f(\mathbf{x}-\theta)$ باشد که در آن f معلوم و θ پارامتر مکان مجهول است. اگر گروه تبدیلات زیر را در نظر بگیریم:

$$\mathcal{G}=\{g\mid g(\mathbf{x})=(x_1+a,\dots,x_n+a)=\mathbf{x}+a,a\in\mathbb{R}\}$$

برای اینکه خانواده توزیع ها تحت $\mathcal{G}$ پایا باشد ، بایستی $\bar{\mathcal{G}}$ به صورت زیر تعریف شود:

$$\bar{\mathcal{G}}=\{\bar{g}\mid\bar{g}(\theta)=\theta+a,a\in\mathbb{R}\}$$

حال اگر مساله برآورد $h(\theta)=\theta$ مورد نظر باشد ، از $\tilde{g}(h(\theta))=h(\bar{g}(\theta))$ خواهیم داشت که:

$$\tilde{\mathcal{G}}=\{\tilde{g}\mid\tilde{g}(\delta)=\delta+a,a\in\mathbb{R}\}$$

مساله ما پایا خواهد بود اگر تابع زیان پایا باشد:

$$L(\theta,\delta)=L(\bar{g}(\theta),\tilde{g}(\delta))=L(\theta+a,\delta+a)\Leftrightarrow L(\theta,\delta)=\rho(\delta-\theta)$$

در این حالت برآوردگر همپایا به فرم $\delta(\mathbf{X}+a)=\delta(\mathbf{X})+a$ است که به آن برآوردگر همپایای مکان می گویند.

قضیه : اگر δ_0 هر برآوردگر همپایای مکان باشد ، آنگاه شرط لازم و کافی برای آنکه برآوردگر δ برای θ همپایا باشد آنست که $\delta(\mathbf{X})=\delta_0(\mathbf{X})+U(\mathbf{X})$ که در آن $U(\mathbf{X})$ هر تابعی است که در شرط زیر صدق کند: $(\forall a,\mathbf{X})\quad U(\mathbf{X}+a)=U(\mathbf{X})$.

قضیه : یک تابع U در شرط $U(\mathbf{X}+a)=U(\mathbf{X})$ صدق می کند اگر و فقط اگر برای $n\geq 2$ تابعی از تفاضلات $(i=1,\dots,n-1)\quad y_i=x_i-x_n$ باشد و برای $n=1$ تابعی ثابت باشد.

قضیه : فرض کنید X دارای چگالی $f_\theta(\mathbf{x})=f(\mathbf{x}-\theta)$ باشد و $(i=1,\dots,n-1)\quad y_i=x_i-x_n$ و $\mathbf{Y}=(Y_1,\dots,Y_{n-1})$. همچنین تابع زیان به فرم $L(\theta,\delta)=\rho(\delta-\theta)$ و δ_0 یک برآوردگر همپایای مکان برای θ با مخاطره متناهی باشد. فرض کنید برای هر $\mathbf{y}$ عددی مانند $v(\mathbf{y})=v^*(\mathbf{y})$ موجود باشد که مقدار $E_{\theta=0}[\rho(\delta_0(\mathbf{X})-v(\mathbf{y}))\mid\mathbf{y}]$ را مینیمم کند، در این صورت یک برآوردگر MRE مکان برای θ عبارتست از : $\delta^*(\mathbf{X})=\delta_0(\mathbf{X})-v^*(\mathbf{y})$.

نتیجه: اگر تابع زیان به فرم $L(\theta,\delta)=(\delta-\theta)^2$ باشد آنگاه $\delta^*(\mathbf{X})=\delta_0(\mathbf{X})-E_{\theta=0}[\delta_0(\mathbf{X})\,|\,\mathbf{y}]$ و اگر $L(\theta,\delta)=|\delta-\theta|$ آنگاه $\delta^*(\mathbf{X})=\delta_0(\mathbf{X})-Med_{\theta=0}[\delta_0(\mathbf{X})\,|\,\mathbf{y}]$.

برآورد گر پیتمن: اگر تابع زیان به فرم درجه دوم خطا باشد ، می توان برآورد گر MRE مکان را از رابطه زیر به دست آورد:

$$\delta^*(\mathbf{X})=\frac{\int_{-\infty}^{+\infty}uf(x_1-u,\ldots,x_n-u)du}{\int_{-\infty}^{+\infty}f(x_1-u,\ldots,x_n-u)du}$$

قضیه : اگر تابع زیان به فرم درجه دوم خطا باشد در این صورت برای پارامتر مکان θ داریم:

الف – اگر $\delta(\mathbf{X})$ هر برآورد گر همپایای مکان با اریبی ثابت b باشد آنگاه $\delta(\mathbf{X})-b$ یک برآورد گر همپایای مکان نااریب است و مخاطره کمتری نسبت به $\delta(\mathbf{X})$ دارد.

ب- برآورد گر یکتای MRE ، یک برآوردگر نا اریب است.

ج – اگر برآوردگر $UMVU$ موجود و همپایا باشد ، آنگاه یک برآورد MRE است.

مخاطره نااریب: فرض کنید تابع زیان به فرم $L(\theta,\delta)$ باشد. گوییم برآورد گر δ برای θ مخاطره نااریب است هرگاه

$$E_\theta[L(\theta,\delta)]\le E_\theta[L(\theta',\delta)]\qquad(\forall\theta'\neq\theta)$$

MRE در خانواده مقیاسی توزیع ها:

فرض کنید X دارای چگالی $f_\sigma(\mathbf{x})=\frac{1}{\sigma^n}f(\frac{x_1}{\sigma},\ldots,\frac{x_n}{\sigma})$ باشد که در آن f معلوم و σ پارامتر مقیاس مجهول است. اگر گروه تبدیلات زیر را در نظر بگیریم:

$$\mathcal{G}=\{g\,|\,g(\mathbf{x})=(cx_1,\ldots,cx_n)=c\mathbf{x},c>0\}$$

برای اینکه خانواده توزیع ها تحت $\mathcal{G}$ پایا باشد ، بایستی $\bar{\mathcal{G}}$ به صورت زیر تعریف شود:

$$\bar{\mathcal{G}}=\{\bar{g}\,|\,\bar{g}(\sigma)=c\sigma,c>0\}$$

حال اگر مساله برآورد $h(\sigma)=\sigma^r$ مورد نظر باشد ، از $\tilde{g}(h(\sigma))=h(\bar{g}(\sigma))$ خواهیم داشت که:

$$\tilde{\mathcal{G}}=\{\tilde{g}\,|\,\tilde{g}(\delta)=c^r\delta,c>0\}$$

مساله ما پایا خواهد بود اگر تابع زیان پایا باشد:

$$L(\sigma,\delta)=L(\bar{g}(\sigma),\tilde{g}(\delta))=L(c\sigma,c^r\delta)\Leftrightarrow L(\sigma,\delta)=\gamma(\frac{\delta}{\sigma^r})$$

در این حالت برآورد گر همپایا به فرم $\delta(c\mathbf{X})=c^r\delta(\mathbf{X})$ است که به آن برآوردگر همپایای مقیاس می گویند.

قضیه : اگر δ_0 هر برآوردگر همپایای σ^r باشد ، آنگاه شرط لازم و کافی برای آنکه برآوردگر δ برای σ^r همپایا باشد آنست که $\delta(\mathbf{X})=\delta_0(\mathbf{X})/U(\mathbf{X})$ که در آن $U(\mathbf{X})$ هر تابعی است که در شرط زیر صدق کند: $U(a\mathbf{X})=U(\mathbf{X})\qquad(\forall a,\mathbf{X})$.

قضیه : یک تابع U در شرط $U(c\mathbf{X})=U(\mathbf{X})$ صدق می کند اگر و فقط اگر برای $n\ge 2$ تابعی از تقسیمات $z_i=x_i/x_n\quad(i=1,\ldots,n-1),z_n=x_n/|x_n|$ باشد و برای $n=1$ تابعی ثابت باشد.

قضیه : فرض کنید X دارای چگالی توام $f_\sigma(\mathbf{x})=\frac{1}{\sigma^n}f(\frac{\mathbf{x}}{\sigma})$ باشد و z_i ها طبق تعریف فوق باشند و $\mathbf{Z}=(Z_1,\dots,Z_n)$. همچنین تابع زیان به فرم $L(\sigma,\delta)=\gamma(\delta/\sigma^r)$ و δ_0 یک برآوردگر هم‌پایای مقیاس برای σ^r با مخاطره متناهی باشد. فرض کنید برای هر $\mathbf{z}$ عددی مانند $w(\mathbf{z})=w^*(\mathbf{z})$ موجود باشد که مقدار $E_{\sigma=1}[\gamma(\delta_0(\mathbf{X})/w(\mathbf{z}))\,|\,\mathbf{z}]$ را مینیمم کند، در این صورت یک برآوردگر MRE مقیاس برای σ^r عبارتست از : $\delta^*(\mathbf{X})=\delta_0(\mathbf{X})/w^*(\mathbf{z})$.

نتیجه : الف – اگر $L(\sigma,\delta)=(\frac{\delta}{\sigma^r}-1)^2$ آنگاه MRE پارامتر σ^r عبارتست از:

$$\delta^*(\mathbf{X})=\frac{\delta_0(\mathbf{X})E_{\sigma=1}[\delta_0(\mathbf{X})\,|\,\mathbf{Z}]}{E_{\sigma=1}[\delta_0^2(\mathbf{X})\,|\,\mathbf{Z}]}$$

ب- اگر $L(\sigma,\delta)=|\frac{\delta}{\sigma^r}-1|$ آنگاه MRE پارامتر σ^r عبارتست از $\delta^*(\mathbf{X})=\delta_0(\mathbf{X})/w^*(\mathbf{z})$ که در آن $w^*(\mathbf{z})$ هر میانه – مقیاس $\delta_0(\mathbf{X})\,|\,\mathbf{Z}$ با $\sigma=1$ می باشد.

برآورد گر پیتمن: اگر $L(\sigma,\delta)=(\frac{\delta}{\sigma^r}-1)^2$ آنگاه MRE پارامتر σ^r را می توان از فرمول زیر محاسبه کرد:

$$\delta^*(\mathbf{X})=\frac{\int_0^\infty v^{n+r-1}f(vx_1,\dots,vx_n)dv}{\int_0^\infty v^{n+2r-1}f(vx_1,\dots,vx_n)dv}$$

MRE در خانواده مکان – مقیاسی توزیع ها:

برآوردگر MRE پارامتر مقیاس:

فرض کنید $\mathbf{X}$ دارای چگالی $f_{\theta,\sigma}(\mathbf{x})=\frac{1}{\sigma^n}f(\frac{x_1-\theta}{\sigma},\dots,\frac{x_n-\theta}{\sigma})$ باشد که در آن f معلوم و θ,σ پارامتر مجهول است. اگر گروه تبدیلات زیر را در نظر بگیریم:

$$\mathcal{G}=\{g\,|\,g(\mathbf{x})=(a+bx_1,\dots,a+bx_n)=a+b\mathbf{x}+,a\in\mathbb{R},b>0\}$$

برای اینکه خانواده توزیع ها تحت $\mathcal{G}$ پایا باشد ، بایستی $\bar{\mathcal{G}}$ به صورت زیر تعریف شود:

$$\bar{\mathcal{G}}=\{\bar{g}\,|\,\bar{g}(\theta,\sigma)=(a+b\theta,b\sigma),a\in\mathbb{R},b>0\}$$

حال اگر مساله برآورد $h(\theta,\sigma)=\sigma^r$ مورد نظر باشد ، از $\tilde{g}(h(\theta,\sigma))=h(\bar{g}(\theta,\sigma))$ خواهیم داشت که:

$$\tilde{\mathcal{G}}=\{\tilde{g}\,|\,\tilde{g}(\delta)=b^r\delta,b>0\}$$

مساله ما پایا خواهد بود اگر تابع زیان پایا باشد:

$$L(\theta,\sigma;\delta)=L(\bar{g}(\theta,\sigma);\tilde{g}(\delta))=L(a+b\theta,b\sigma;b^r\delta)\Leftrightarrow L(\theta,\sigma;\delta)=\gamma(\frac{\delta}{\sigma^r})$$

در این حالت برآورد گر هم‌پایا به فرم $\delta(a+b\mathbf{X})=b^r\delta(\mathbf{X})$ است.

حال اگر $(i=1,\dots,n-1)$ $Y_i=X_i-X_n$ و $\mathbf{Y}=(Y_1,\dots,Y_n)$ آنگاه تابع چگالی توام $f_{\mathbf{Y}}$ متعلق به یک خانواده مقیاسی از توزیع ها است لذا می توان مانند قبل برای برآورد σ^r از رابطه $\delta^*(\mathbf{X})=\delta_0(\mathbf{Y})/w^*(\mathbf{z})$ که در آن $\delta_0(\mathbf{Y})$ هر برآورد گر هم‌پایای

مقیاس بر اساس $\mathbf{Y}$ است ؛ یعنی $\delta_{\cdot}(b\mathbf{Y})=b^{r}\delta_{\cdot}(\mathbf{Y})$ یا $\delta_{\cdot}(a+b\mathbf{X})=b^{r}\delta_{\cdot}(\mathbf{X})$. $\mathbf{Z}=(Z_{1},...,Z_{n-1})$ عبارتست از $Z_{i}=Y_{i}/Y_{n-1}$ و $Z_{n-1}=Y_{n-1}/|Y_{n-1}|$ و $w^{*}(\mathbf{z})$ هر مقداری است که امید شرطی زیر را مینیمم کند:

$$E_{\sigma=1}[\gamma\left(\delta_{\cdot}(\mathbf{Y})/w(\mathbf{z})\right)|\,\mathbf{z}]$$

نتیجه : الف – اگر $L(\sigma,\delta)=(\frac{\delta}{\sigma^{r}}-1)^{2}$ آنگاه MRE پارامتر σ^{r} عبارتست از:

$$\delta^{*}(\mathbf{X})=\frac{\delta_{\cdot}(\mathbf{Y})E_{\sigma=1}[\delta_{\cdot}(\mathbf{Y})\,|\,\mathbf{Z}]}{E_{\sigma=1}[\delta_{\cdot}^{2}(\mathbf{Y})\,|\,\mathbf{Z}]}$$

ب- اگر $L(\sigma,\delta)=|\frac{\delta}{\sigma^{r}}-1|$ آنگاه MRE پارامتر σ^{r} عبارتست از $\delta^{*}(\mathbf{X})=\delta_{\cdot}(\mathbf{Y})/w^{*}(\mathbf{z})$ که در آن $w^{*}(\mathbf{z})$ هر میانه – مقیاس $\delta_{\cdot}(\mathbf{Y})\,|\,\mathbf{Z}$ با $\sigma=1$ می باشد.

برآوردگر MRE پارامتر مکان:

اگر گروه تبدیلات زیر را در نظر بگیریم:

$$\mathcal{G}=\{g\,|\,g(\mathbf{x})=(a+bx_{1},...,a+bx_{n})=a+b\mathbf{x}+,a\in\mathbb{R},b>0\}$$

برای اینکه خانواده توزیع ها تحت $\mathcal{G}$ پایا باشد ، بایستی $\bar{\mathcal{G}}$ به صورت زیر تعریف شود:

$$\bar{\mathcal{G}}=\{\bar{g}\,|\,\bar{g}(\theta,\sigma)=(a+b\theta,b\sigma),a\in\mathbb{R},b>0\}$$

حال اگر مساله برآورد $h(\theta,\sigma)=\theta$ مورد نظر باشد ، از $\tilde{g}(h(\theta,\sigma))=h(\bar{g}(\theta,\sigma))$ خواهیم داشت که:

$$\tilde{\mathcal{G}}=\{\tilde{g}\,|\,\tilde{g}(\delta)=a+b\delta,a\in\mathbb{R},b>0\}$$

مساله ما پایا خواهد بود اگر تابع زیان پایا باشد:

$$L(\theta,\sigma;\delta)=L(\bar{g}(\theta,\sigma);\tilde{g}(\delta))=L(a+b\theta,a+b\sigma;b^{r}\delta)\Leftrightarrow L(\theta,\sigma;\delta)=\rho(\frac{\delta-\theta}{\sigma})$$

در این حالت برآورد گر همپایا به فرم $\delta(a+b\mathbf{X})=a+b\delta(\mathbf{X})$ است.

قضیه : فرض کنید $\delta_{\cdot}(\mathbf{X})$ هر برآوردگر همپایای θ باشد ؛ یعنی $\delta_{\cdot}(a+b\mathbf{X})=a+b\delta_{\cdot}(\mathbf{X})$ و $\delta_{1}(\mathbf{X})$ هر برآوردگر همپایای σ باشد یعنی $\delta_{1}(a+b\mathbf{X})=b\delta_{1}(\mathbf{X})$ در این صورت برآوردگر $\delta(\mathbf{X})$ همپایای مکان در خانواده مکان – مقیاس است اگر و فقط اگر $\delta(\mathbf{X})=\delta_{\cdot}(\mathbf{X})-w(\mathbf{z})\delta_{1}(\mathbf{X})$ که در آن $(i=1,...,n-1)$ $Y_{i}=X_{i}-X_{n}$ و $\mathbf{Z}=(Z_{1},...,Z_{n-1})$ عبارتست از $(i=1,...,n-1)$ $Z_{i}=Y_{i}/Y_{n-1}$ و $Z_{n-1}=Y_{n-1}/|Y_{n-1}|$.

نتیجه : برآوردگر MRE پارامتر θ $\delta^{*}(\mathbf{X})=\delta_{\cdot}(\mathbf{X})-w^{*}(\mathbf{z})\delta_{1}(\mathbf{X})$ است که در آن $w^{*}(\mathbf{z})$ هر مقداری است که امید شرطی زیر را مینیمم کند:

$$E_{\theta=0,\sigma=1}[\rho\left(\delta_{\cdot}(\mathbf{X})-w(\mathbf{z})\delta_{1}(\mathbf{X})\right)|\,\mathbf{z}]$$

پس اگر $L(\theta,\sigma;\delta)=(\frac{\delta-\theta}{\sigma})^{2}$ آنگاه MRE پارامتر θ عبارتست از $\delta^{*}(\mathbf{X})=\delta_{\cdot}(\mathbf{X})-w^{*}(\mathbf{z})\delta_{1}(\mathbf{X})$ که در آن

$$w^{*}(\mathbf{Z})=\frac{E_{0,1}[\delta_{\cdot}(\mathbf{X})\delta_{1}(\mathbf{X})\,|\,\mathbf{Z}]}{E_{0,1}[\delta_{1}^{2}(\mathbf{X})\,|\,\mathbf{Z}]}$$