فهرست

1- مقدمه

نظریه برآورد به دو موضوع برآورد نقطه‌ای و برآورد فاصله‌ای می‌پردازد. در برآورد نقطه‌ای یک آماره، تنها یک مقدار عددی بکار می‌رود تا بتوان به کمک آن یک پارامتر بخصوصی از جامعه را برآورد نمود در حالیکه در برآورد فاصله‌ای، یک فاصله مخصوصی معین می‌گردد که مقدار واقعی پارامتر در داخل این فاصله قرار دارد.

2- برآورد نقطه‌ای(Point Estimator)

یک مقدار به خصوصی از یک آماره که برای برآورد یک پارامتر مشخص به کار می‌رود به نام برآورد نقطه ای نامیده می شود. می توان مساله کلی برآورد نقطه‌ای را به شرح ذیل بیان کرد:

متغیر تصادفی X، با تابع چگالی احتمال $f_X(x;\theta)$ مفروض است که در آن پارامتر θ مجهول می‌باشد. یک نمونه تصادفی، $x_n,\ldots\ldots,x_2,x_1$ از این جامعه انتخاب می‌نماییم و بر اساس تابعی از این نمونه تصادفی، $\hat{\theta}=\hat{\theta}(x_1,x_2,\ldots\ldots,x_n)$، مقدار پارامتر θ را برآورد می‌کنیم. در این صورت متغیر تصادفی $\hat{\theta}$ یک برآورد کننده برای پارامتر θ و مقداری که $\hat{\theta}$ می‌گیرد برآورد نقطه‌ای θ نامیده می‌شود.

بدیهی است، برای پارامتر θ، تعداد زیادی برآوردکننده وجود دارد. اما یک برآوردکننده خوب آن است که تا حد ممکن به مقدار واقعی پارامتر نزدیک باشد. در اینجا این سوال مطرح می‌شود که از بین تخمین زننده‌های مختلف برای پارامتر θ کدام بهتر است؟

از آنجایی که تخمین زننده‌ها، متغیرهای تصادفی هستند پس عملکرد آنها را نمی‌توان برای یک مورد خاص مدنظر قرار دارد. بنابراین برآوردکننده‌ها باید براساس عملکردشان در بلندمدت مورد ارزیابی قرار گیرند.

3- ارزیابی یک تخمین زننده یا مقایسه برآوردکننده ها

در برآورد پارامتر θ که بر اساس مقادیر نمونه تعیین می شود، در ابتدا تابع ضرر (زیان) به صورت ذیل تعریف می کنیم:

$$\ell(\hat{\theta},\theta)=(\hat{\theta}-\theta)^2$$

از آنجایی که تابع زیان یک متغیر تصادفی است بنابراین منطقی بنظر می‌رسد از امید ریاضی تابع

زیان تحت عنوان تابع ریسک برای نیل به ارزیابی بهتر استفاده شود. تابع ریسک عبارتست از:

$$R(\hat{\theta},\theta)=MSE=E[\ell(\hat{\theta},\theta)]=E(\hat{\theta}-\theta)^2$$

3-1- مقایسه دو برآوردکننده از طریق سنجش کارآیی نسبی آنها

اگر برآورد کننده $\hat{\theta}_1$ دارای میانگین مربع خطای $E(\hat{\theta}_1-\theta)^2$ (که آنرا با MSE1 نشان می‌دهند) و برآوردکننده $\hat{\theta}_2$ دارای میانگین مربع خطای $E(\hat{\theta}_2-\theta)^2$ (که آنرا با MSE2 نشان می‌دهند) باشد در این صورت کارایی $\hat{\theta}_2$ نسبت به $\hat{\theta}_1$ بصورت $\frac{MSE_1}{MSE_2}$ یا $\frac{E(\hat{\theta}_1-\theta)^2}{E(\hat{\theta}_2-\theta)^2}$ تعریف می‌شود. اگر کارایی $\hat{\theta}_2$ نسبت به $\hat{\theta}_1$ کمتر از یک باشد، $\hat{\theta}_1$ برآورد کننده ای بهتر از $\hat{\theta}_2$ برای θ به شمار می آید. در غیر این صورت $\hat{\theta}_2$ برآوردکننده ای بهتر از $\hat{\theta}_1$ برای θ است.

3-2- تعریف تخمین زننده بهینه

تخمین زننده بهینه به تخمین زننده‌ای گفته می‌شود که MSE مربوط به آن، از MSE سایر تخمین زننده ها به ازای تمام مقادیر θ کمتر باشد. به عنوان مثال $\overline{X}$ یک تخمین زننده بهینه μ مربوط به توزیع نرمال می باشد.

3-3- برآوردکننده های نااریب

$\hat{\theta}$ یک برآوردکننده نااریب برای θ است اگر امید ریاضی آن با θ مساوی شود. یعنی اگر به ازای تمام مقادیر θ داشته باشیم:

$$E(\hat{\theta})=\theta$$

اگر یک برآوردکننده ناریب نباشد، اریب نامیده می شود و مقدار $b=E(\hat{\theta})-\theta$ به عنوان اریبی شناخته می شود. در اینصورت، اگر $b>0$ باشد اریبی مثبت و اگر $b<0$ باشد اریبی منفی است. در اینصورت میانگین مربعات خطا بر حسب مقدار اریبی عبارتست از:

$$MSE(\hat{\theta})=E(\hat{\theta}-\theta)^2=Var(\hat{\theta})+b^2$$

نکته: اگر متغیر تصادفی T یک برآوردکننده نااریب برای پارامتر مجهول θ باشد و $Var(T)\neq 0$ در این صورت T^2 همواره برای θ^2 اریب است.

مثال: دو برآوردکننده $\hat{\theta}_1$ و $\hat{\theta}_2$ برای پارامتر θ به شرح ذیل بیان شده است:

$$\begin{cases} E(\hat{\theta}_1)=\theta \\ Var(\hat{\theta}_1)=7 \end{cases} \qquad \begin{cases} E(\hat{\theta}_2)=\frac{2}{3}\theta \\ Var(\hat{\theta}_2)=3 \end{cases}$$

بر حسب مقادیر مختلف θ این دو تخمین زننده را مقایسه کنید.

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + b^2$$

$$MSE(\hat{\theta}_1) = 7 + \left[E(\hat{\theta}_1) - \theta_1\right]^2 = 7 + (0)^2 = 7$$

$$MSE(\hat{\theta}_2) = 3 + \left[E(\hat{\theta}_2) - \theta_2\right]^2 = 3 + \left(\frac{2}{3}\theta - \theta\right)^2 = 3 + \frac{1}{9}\theta^2$$

اگر $\hat{\theta}_2$ برآورد کننده بهتری برای θ باشد آنگاه:

$$MSE(\hat{\theta}_1) > MSE(\hat{\theta}_2)$$

$$7 > 3 + \frac{\theta^2}{9} \rightarrow -6 < \theta < 6$$

مثال 1: متغیر تصادفی X دارای توزیع نرمال با میانگین مجهول μ و واریانس مجهول σ^2 است. یک نمونه تصادفی 9 تایی می‌گیریم و 2 برآورد کننده نقطه‌ای $T_1 = 0.1\sum_{i=1}^{9}(x_i - \bar{x})^2$ و $T_2 = 0.125\sum_{i=1}^{9}(x_i - \bar{x})^2$ را برای σ^2 تعریف می‌کنیم اگر $Var(T_1) = 0.16\sigma^4$ و $Var(T_2) = 0.25\sigma^4$ باشد. کارایی T_1 نسبت به T_2 برابر است با:

$$\frac{\text{کارایی } T_1}{\text{کارایی } T_2} = \frac{MSE_2}{MSE_1} = \frac{Var(T_2) + (E(T_2) - \sigma^2)^2}{Var(T_1) + (E(T_1) - \sigma^2)^2} = \frac{0.25\sigma^4}{0.16\sigma^4 + 0.04\sigma^4} = 1.25$$

$$E(T_1) = 0.1 \times 8 \times \sigma^2 = 0.8\sigma^2$$

$$E(T_2) = 0.125 \times 8\sigma^2 = \sigma^2$$

مثال 2: اگر X_1 و X_2 و X_3 یک نمونه تصادفی از جامعه نرمال با میانگین μ و واریانس σ^2 باشند، کارائی نسبی برآورد کننده $\bar{X}$ (میانگین نمونه) نسبت به برآوردکننده $\frac{x_1 + 2x_2 + x_3}{4}$ چقدر است؟

$$\frac{\text{کارایی } \bar{x}}{\text{کارایی } y} = \frac{MSE(y)}{MSE(\bar{x})} = \frac{Var(y)}{Var(\bar{x})} = \frac{\frac{1}{16} \times 6\sigma^2}{\frac{\sigma^2}{3}} = \frac{9}{8}$$

مثال 3: فرض کنید T_1 و T_2 برآوردکننده های θ هستند و بصورت زیر بیان می‌شوند:

$$\begin{cases} E(T_1)=\theta \\ Var(T_1)=4 \end{cases} \qquad \begin{cases} E(T_2)>\theta \\ Var(T_2)=2 \end{cases}$$

و $[E(T_2)-\theta]^2=7$ است. کدام برآوردکننده بهتری برای θ است؟

$$MSE_{T_1}=Var(T_1)+b^2=4+0=4$$

$$MSE_{T_2}=Var(T_2)+[E(T_2)-\theta]^2=2+7=9$$

چون $\frac{MSE_{T_1}}{MSE_{T_2}}<1$ است لذا T_1 برآوردکننده بهتری برای θ است.

مثال 4: فرض کنید T بر اساس یک نمونه تصادفی n تایی، برآوردکننده‌ای برای θ باشد. اگر $E(x_i)=\theta$ و $T=\sum a_i x_i$ باشد. a_i ها چه محدودیتی باید داشته باشند تا T یک برآوردکننده ناریب θ شود؟

$$T=\sum a_i x_i \rightarrow E(T)=E\left[\sum a_i x_i\right]=\sum[E(a_i x_i)]=\sum[a_i E(x_i)]$$

چون T یک برآوردکننده ناریب θ است لذا:

$$\theta=\sum(a_i\theta)=\theta=\theta\sum a_i \rightarrow \sum a_i=1$$

مثال 5: فرض کنید $X\sim N(\mu,\sigma^2)$ باشد. اگر $\hat{\theta}_1$ و $\hat{\theta}_2$ دو تخمین زننده برای پارامتر σ^2 باشند، آنها را مقایسه کنید.

$$\hat{\theta}_1=\frac{\sum(x_i-\bar{x})^2}{n-1} \qquad , \hat{\theta}_2=\frac{\sum(x_i-\bar{x})^2}{n+1}$$

$$E(\hat{\theta}_1)=E(S^2)=\sigma^2 \quad , \quad Var(\hat{\theta}_1)=Var(S^2)=\frac{2\sigma^4}{n-1} \quad , \quad MSE(\hat{\theta}_1)=\frac{2\sigma^4}{n-1}$$

$$E(\hat{\theta}_2)=E\left[\frac{(n-1)S^2}{(n+1)}\right]=\frac{n-1}{n+1}\sigma^2$$

$$Var(\hat{\theta}_2)=Var\left[\frac{(n-1)}{(n+1)}S^2\right]=\frac{2(n-1)}{(n+1)^2}\sigma^4$$

$$MSE(\hat{\theta}_2)=\frac{2(n-1)}{(n+1)^2}\sigma^4+\left[\frac{n-1}{n+1}\sigma^2-\sigma^2\right]^2=\frac{2\sigma^4}{n+1}$$

$MSE(\hat{\theta}_1)>MSE(\hat{\theta}_2)\Rightarrow$ $\hat{\theta}_2$ برآورد کننده بهتر از $\hat{\theta}_1$ برای پارامتر σ^2 است

3-4- برآوردکننده های سازگار

خاصیت نااریبی بر حسب تکرار آزمایش‌ها بیان شده است اما خاصیت سازگاری مربوط به رفتار یک تخمین زننده در یک آزمایش است وقتیکه حجم نمونه اجازه یابد بسیار بزرگ شود.

تعریف: گفته می شود برآورد کننده $\hat{\theta}_{(n)}$ یک برآورد کننده سازگار است هرگاه:

1- $\hat{\theta}_{(n)}$ نااریب باشد.

2-با بزرگ شدن اندازه نمونه (n) داشته باشیم:

$$\lim_{n\to\infty}\hat{\theta}_{(n)}=\theta \qquad , \qquad \lim_{n\to\infty}\operatorname{Var}\left(\hat{\theta}_{(n)}\right)=0$$

نماد $\hat{\theta}_{(n)}$ به کار می رود تا نشان دهد که ممکن است برآوردکننده تابعی از نمونه تصادفی باشد.

بطور کلی برآوردکننده‌هایی که تابع ریسک آنها، با بزرگ شدن اندازه نمونه به سمت صفر میل می‌گراید، برآوردکننده های سازگارند.

اگر این برآوردکننده ها اریب باشند، مقدار اریبی آنها نیز به سمت صفر میل می کند.

$$\lim_{n\to\infty}MSE_\theta(\theta_n)=0$$

مثال 6: نمونه تصادفی $x_1,\dots\dots,x_n$ را از متغیر تصادفی دلخواه X با میانگین μ واریانس σ^2 در نظر می گیریم. فرض کنید:

$$\bar{x}_n=\frac{1}{n}\sum x_i; \qquad U_n=\bar{x}_n+1$$

نشان دهید که $\overline{X}_n$ یک برآوردکننده سازگار و U_n یک برآوردکننده ناسازگار برای μ می باشد.

$\overline{X}_n$ یک برآوردکننده سازگار نااریب برای μ است.

$$E(\overline{X}_n)=\mu\ ,\ \lim_{n\to\infty}\operatorname{Var}(\overline{X}_n)=\lim_{n\to\infty}\frac{\sigma^2}{n}=0$$

U_n یک برآوردکننده ناسازگار است.

$$E(U_n)=E(\overline{X}_n+1)=\mu+1$$

نکته: سازگاری یک خاصیت مجانبی است یعنی خاصیت حدی یک برآوردکننده است. به عبارت دیگر وقتی n به حد کافی بزرگ است می توانیم عملاً مطمئن باشیم که خطایی که با یک برآوردکننده سازگار صورت می گیرد از هر ثابت مثبت مفروضی کمتر خواهد بود.

نکته: شرط کافی (نه لازم) برای اینکه آماده $\hat{\theta}$ یک برآوردکننده سازگار پارامتر θ باشد این است که:

1) $\hat{\theta}$ نااریب باشد.

2)وقتی $n \to \infty$ ، $\operatorname{Var}(\hat{\theta}) \to 0$

این شرط، شرط کافی است و شرط لازم نیست. لذا یک برآوردکننده می تواند سازگار باشد بدون اینکه نااریب باشد.

مثال 7: نشان دهید برآوردکننده $\frac{x+1}{n+2}$ یک برآوردکننده سازگار پارامتر θ جامعه دو جمله ای است.

نکته: یک برآوردکننده اریب تنها وقتی می تواند سازگار باشد که بطور مجانبی نااریب باشد یعنی وقتی $n \to \infty$ نااریب باشد.

مثال 8: برآوردکننده مینیماکس پارامتر θ ی دو جمله ای یعنی $\frac{X+\frac{1}{2}\sqrt{n}}{n+\sqrt{n}}$ مجانباً نااریب است.

$$E\left(\frac{x+\frac{1}{2}\sqrt{n}}{n+\sqrt{n}}\right)=\frac{E(x)+\frac{1}{2}\sqrt{n}}{n+\sqrt{n}}+\frac{n\theta+\frac{1}{2}\sqrt{n}}{n+\sqrt{n}}\neq\theta$$

$$\lim_{n\to\infty} E\left(\frac{x+\frac{1}{2}\sqrt{n}}{n+\sqrt{n}}\right)=\lim_{n\to\infty}\frac{n\theta+\frac{1}{2}\sqrt{n}}{n+\sqrt{n}}=\theta$$

مثال 9: نشان دهید که واریانس نمونه ای S^2، برآوردکننده سازگار σ^2 برای یک نمونه تصادفی از جامعه های نرمال است.

1) $E(S^2)=\sigma^2$ نااریب است S^2

2) $\operatorname{Var}(S^2)=\frac{2\sigma^4}{n-1}$

$$\lim_{n\to\infty}\operatorname{Var}(S^2)=\lim_{n\to\infty}\frac{2\sigma^4}{n-1}=0$$

مثال 10: نشان دهید که $\bar{x}^2$ یک برآوردکننده مجانباً نااریب μ^2 است.

$$E(\bar{x}^2) = Var(\bar{x}) + [E(\bar{x})]^2$$

$$E(\bar{x}^2) = \frac{\sigma^2}{n} + \mu^2$$

$$\lim_{n\to\infty} E(\bar{x}^2) = \lim_{n\to\infty}\left(\frac{\sigma^2}{n} + \mu^2\right) = \mu^2$$

پس $\bar{x}^2$ یک برآوردکننده مجانباً نااریب است.

5-3- تخمین زننده های نااریب کارا

اگر برآوردکننده‌های پارامتر θ متعلق به طبقه برآوردکننده نااریب باشند تابع ریسک، همان واریانس برآوردکننده می‌شود. در صورت وجود، برآوردکننده بهینه، $\hat{\theta}_0$، در میان این طبقه برآوردکننده ای است که به ازای تمام مقادیر θ واریانس آن از واریانس هیچ برآوردکننده نااریب دیگر تجاوز نکند. اگر چنین تخمین زننده ای وجود داشته باشد تخمین زننده $\hat{\theta}_0$ به تخمین زننده نااریب با حداقل واریانس θ معروف بوده و نیز در این طبقه از همه کاراتر است.

برای جستجو برای یک چنین برآورد کننده‌ای، تعیین یک حد پائین برای واریانس تمام برآوردکننده های نااریب مفید خواهد بود. یک چنین حدی بوسیله یک نامساوی مشهور به نامساوی کرایمر- رائو ارائه می شود.

فرض کنید $\hat{\theta}$، یک برآورد کننده نااریب برای پارامتر θ از جامعه‌ای با تابع چگالی $f_x(x;\theta)$باشد. تحت شرایط بسیار کلی $\hat{\theta}$ در نامساوی ذیل صدق می‌کند.

$$Var(\hat{\theta}) \geq \frac{1}{nE\left[\left(\frac{\partial Lnf(x;\theta)}{\partial\theta}\right)^2\right]}$$

است. اگر X گسسته باشد، $f_x(x;\theta)$ را برداشته و به جای آن از $P(x,\theta)$ استفاده می کنیم.

متاسفانه همواره برآوردکننده‌ای که عملاً دارای حد پائین کریمر- رائو باشد وجود ندارد. اما در بسیاری از موارد می‌توان نشان داد که برآوردکننده‌ای مانند $\hat{\theta}_0$ دارای چنین حدی است در این صورت $\hat{\theta}_0$ باید برآوردکننده بهینه نااریب باشد. به عبارت دیگر حد پائین کریمر- رائو متعلق به

تخمین زننده نااریب بهینه است.

قضیه: اگر $\hat{\theta}$ یک برآوردکننده نااریب θ باشد و $\mathrm{Var}(\hat{\theta}) = \dfrac{1}{nE\left[\left(\dfrac{\partial \mathrm{Lnf}(x;\theta)}{\partial\theta}\right)^2\right]}$ باشد، آنگاه $\hat{\theta}$ یک برآوردکننده نااریب با کمترین واریانس از θ است.

مثال11: نشان دهید $\overline{X}$ یک برآوردکننده نااریب با کمترین واریانس برای میانگین جامعه نرمال است.

$$X \sim N(\mu, \sigma^2)$$

$$f_X(x;\mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \qquad ; \qquad -\infty < x < \infty$$

$$\mathrm{Ln}\, f_X(x;\mu) = \mathrm{Ln}\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2\sigma^2}(x-\mu)^2$$

$$\frac{d\,\mathrm{Ln}\, f_X(x;\mu)}{d\mu} = \frac{x-\mu}{\sigma^2}$$

$$E\left\{\frac{d\,\mathrm{Ln}\, f_X(x;\mu)}{d\mu}\right\}^2 = \frac{E(x-\mu)^2}{\sigma^4} = \frac{\sigma^2}{\sigma^4} = \frac{1}{\sigma^2}$$

$$\frac{1}{nE\left\{\dfrac{d\,\mathrm{Ln}\, f_X(x;\mu)}{d\mu}\right\}^2} = \frac{\sigma^2}{n} \qquad (1)$$

$$\mathrm{Var}(\overline{x}) = \frac{\sigma^2}{n} \qquad (2)$$

$$(1) = (2)$$

لذا واریانس میانگین نمونه به حد پائین کریمر-رائو می رسد. بنابراین $\overline{X}$ برآوردکننده نااریب کمترین واریانس برای میانگین توزیع نرمال با واریانس معلوم است.

4- روش‌های برآورد نقطه‌ای

4-1- روش گشتاورها

فرض کنید $x_1,.....,x_n$ یک نمونه تصادفی n تایی از یک توزیع دلخواه باشد. در اینصورت k امین گشتاور نمونه M'_k حول مبدا امین گشتاور توزیع حول مبدا بترتیب عبارتست از $M'_k = \frac{\sum_{i=1}^{n} X_i^k}{n}$ و $E(X^k)$ که گشتاور توزیع حول مبدا، عموماً توابعی از p پارامتر مجهول است. در روش گشتاورها با برابر قرار دادن p گشتاور اول توزیع با گشتاورهای نظیر نمونه

$$M'_1 = E(x) \qquad \bar{x} = E(x)$$
$$M'_2 = E(x^2) \qquad \frac{\sum x_i^2}{n} = E(x^2)$$
$$\vdots \qquad \vdots$$
$$\vdots \qquad \vdots$$
$$M'_p = E(x^p) \qquad \frac{\sum x_i^p}{n} = E(x^p)$$

و حل دستگاه حاصل پارامترهای مجهول جامعه بدست می‌آید. بطور خلاصه: این روش عبارتست از مساوی قرار دادن گشتاورهای جامعه با گشتاورهای نمونه و بعد حل دستگاه معادلات حاصل.

مثال12: برآورد کننده های μ و σ^2 یک جامعه نرمال با میانگین مجهول μ و واریانس مجهول σ^2 را به روش گشتاورها تعیین کنید.

برای یک متغیر تصادفی با توزیع نرمال داریم:

$$E(x) = \mu$$
$$E(x^2) = \sigma^2 + \mu^2$$

در این صورت با مساوی قرار دادن دو گشتاور اول نمونه با دو گشتاور اول توزیع داریم:

$$\bar{x} = \mu$$
$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 = \sigma^2 + \mu^2$$

می انجامد. از حل توام این معادلات، برآورد کننده های روش گشتاورها عبارتند از:

$$\hat{\mu} = \bar{x}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$$

مثال13: اگر $x_1,\ldots,x_n$ مقادیر یک نمونه تصادفی به اندازه n از جامعه‌ای با چگالی

$f(x;\theta) = \begin{cases} \frac{2(\theta - x)}{\theta^2} & ; \ 0 < x < \theta \\ 0 & \text{سایر مقادیر} \end{cases}$ باشد. برآورد کننده ای برای θ به روش گشتاورها پیدا کنید.

$$E(x) = \int_0^\theta x \frac{2(\theta - x)}{\theta^2} dx = \frac{2}{\theta^2} \int_0^\theta (\theta x - x^2) dx = \frac{2}{\theta^2} \left[\frac{1}{2}\theta x^2 - \frac{1}{3} x^3 \right]_0^\theta = \frac{2}{\theta^2} \left(\frac{1}{2}\theta^3 - \frac{1}{3}\theta^3 \right) =$$

$$\frac{2}{\theta^2} \cdot \frac{1}{6}\theta^3 = \frac{1}{3}\theta$$

$$M'_1 = E(x)$$

$$\bar{x} = \frac{1}{3}\theta \rightarrow \hat{\theta} = 3\bar{x}$$

مثال14: اگر $x_1,\ldots,x_n$ مقادیر یک نمونه تصادفی به اندازه n از جامعه هندسی باشد. برآورد کننده پارامتر θ ی توزیع را به روش گشتاورها بدست آورید.

$$f_x(x;\theta) = \theta(1-\theta)^{x-1}$$

$$E(x) = \frac{1}{\theta} \qquad \bar{x} = \frac{1}{\theta} \rightarrow \theta = \frac{1}{x}$$

مثال15: نمونه تصادفی به اندازه n از جامعه گاما داریم. برای برآورد کردن پارامترهای α و β روش گشتاورها را بکار ببرید.

$$\left.\begin{aligned} &\bar{x} = \alpha\beta \\ &\frac{\sum x_i^2}{n} = \alpha\beta^2 + \bar{x}^2 \end{aligned}\right\} \rightarrow \hat{\alpha} = \frac{n\bar{x}^2}{\sum (x_i - \bar{x})^2} \quad , \quad \hat{\beta} = \frac{\sum (x_i - \bar{x})^2}{n\bar{x}}$$

مثال16: اگر $x_1,\ldots,x_n$ مقادیر یک نمونه تصادفی به اندازه n از جامعه بتا با $\beta = 1$ باشد برآوردی برای پارامتر α به روش گشتاورها بدست آورید.

$$f_x(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

$$E(x)=\frac{\alpha}{\alpha+\beta} \qquad \bar{x}=\frac{\alpha}{\alpha+1}\rightarrow\hat{\alpha}=\frac{\bar{x}}{1-\bar{x}}$$

مثال17: تابع چگالی زیر داده شده است که در آن θ پارامتر مجهول می باشد. θ را به روش گشتاورها تخمن بزنید.

$$f_x(x)=\theta x^{\theta-1} \quad ; \quad 0<x<1 \quad ; \ 0<\theta<\infty$$

$$E(x)=\int_0^1 x\theta x^{\theta-1}dx=\theta\int_0^1 x^{\theta}dx=\frac{\theta}{\theta+1}\left[x^{\theta+1}\right]_0^1=\frac{\theta}{\theta+1}$$

$$\frac{\theta}{\theta+1}=\bar{x}\rightarrow\hat{\theta}=\frac{\bar{x}}{1-\bar{x}}$$

4-2- روش حداکثر درستنمایی (روش MLE)

این روش عبارتست از حداکثر کردن درست نمایی احتمال بدست آوردن یک مجموعه از مقادیر نمونه تصادفی. فرض کنید $x_1,x_2,.....,x_n$ یک نمونه به حجم n از یک جامعه پیوسته با چگالی احتمال $f_x(x,\theta)$ یا گسسته با احتمال $P_x(x,\theta)$ باشد که θ یک پارامتر مجهول است و باید برآورد شود. در حقیقت می خواهیم بجای برآورنده θ، آن تابعی از مشاهدات نمونه ای را قرار دهیم که احتمال نمونه به ازای آن حداکثر شود. تابع احتمال L را بصورت:

$$L=P(x_1,x_2,....,x_n|\theta)=\prod_{i=1}^{n}P(x_i=x_i)$$

$$L=f_x(x_1,x_2,....,x_n|\theta)=\prod_{i=1}^{n}f_x(x_i,\theta)$$

تعریف می کنیم که همان احتمال توام برای مشاهدات نمونه است. تخمین θ باید به گونه‌ای صورت پذیرد که بیشترین احتمال برای مشاهده مقادیر $x_1,x_2,.....,x_n$ حاصل شود. به عبارت دیگر تخمین زننده بیشترین احتمال $\hat{\theta}$ مقداری خواهد بود که به ازای آن تابع L حداکثر شود. به تابع L، تابع درست نمایی می‌گویند و $\hat{\theta}$ حاصل از این روش را با نماد $\hat{\theta}=MLE(\theta)$ نمایش می دهیم.

توجه: اغلب در حداکثر کردن تابع $f(x_1,....,x_n|\theta)$ راحت‌تر است که Ln آن حداکثر شود زیرا

Ln(L) یک تابع صعودی یکنواخت از L است و θ که L را ماکزیمم می‌کند Ln(L) را نیز ماکزیمم می کند.

مثال18: اگر $x_1, x_2, \ldots, x_n$ مقادیر یک نمونه تصادفی به اندازه n از یک جامعه نرمال $N(\mu, \sigma^2)$ باشد، برآورد کننده ای برای μ و σ بیابید.

$$L = f(x_1, \ldots, x_n|\theta) = \prod_{i=1}^{n} f(x_i|\mu,\sigma) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} =$$

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

$$LnL = -n\,Ln\,\sigma\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial\,Ln\,L}{\partial\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0 \rightarrow \sum_{i=1}^{n}(x_i-\mu) = 0 \rightarrow \hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

$$\frac{\partial LnL}{\partial\delta} = -\frac{n}{\delta} + \frac{1}{\delta^3}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

مثال19: اگر x_1 . .. x_n مقادیر یک نمونه تصادفی به اندازه n از جامعه برنولی باشد. MLE پارامتر برنولی را بدست آورید.

$$p_x(x_i, p) = p^{x_i}(1-p)^{1-x_i}, x_i = 0,1$$

$$L = p^{\sum x_i}(1-p)^{n-\sum x_i}$$

$$LnL = (Lnp)\left(\sum_{i=1}^{n} x_i\right) + [Ln(1-p)]\left(n - \sum_{i=1}^{n} x_i\right)$$

$$\frac{dLnL}{dp} = \frac{\sum_{i=1}^{n} x_i}{p} - n - \frac{\sum_{i=1}^{n} x_i}{1-p} = 0 \Rightarrow \frac{\sum x_i}{p} = \frac{n-\sum x_i}{1-p} \Rightarrow \sum_{i=1}^{n} x_1 = np \Rightarrow p = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

مثال20: اگر x_1 و... وx_2 نمونه ای به اندازه n از جامعه دو جمله ای با پارامتر $p = \theta$ باشد. MLE

پارامتر دوجمله‌ای را بدست آورید.

راه حل اول:

$$L(\theta)=f(x;\theta)=\binom{n}{n}\theta^{x}(1-\theta)^{n-x}$$

$$\frac{\partial f}{\partial \theta}=\binom{n}{x}\left[x\theta^{x-1}(1-\theta)^{n-x}-(n-x)\theta^{x}(1-\theta)^{n-x-1}\right]=0$$

$$f_X(x;\theta)\left[x\theta^{-1}-(n-x)(1-\theta)^{-1}\right]=0$$

$$\frac{x}{\theta}-\frac{n-x}{1-\theta}=0\Rightarrow x=n\theta\Rightarrow\hat{\theta}=\frac{x}{n}$$

راه حل دوم:

$$\operatorname{Ln}L(\theta)=\operatorname{Ln}\binom{n}{x}+x\operatorname{Ln}\theta+(n-x)\operatorname{Ln}(1-\theta)$$

$$\frac{d\operatorname{Ln}L(\theta)}{d\theta}=\frac{x}{\theta}-\frac{n-x}{1-\theta}=0\Rightarrow\hat{\theta}=\frac{x}{n}$$

خواص MLE

1- با افزایش n توزیع MLE به توزیع نرمال میل می‌کند و درزمانیکه $MLE, n\to\infty$ توزیع نرمال دارد.

2-در زمانیکه $n\to\infty$ ، MLE سازگار است و به تبع نااریب است.

3- اگر $\hat{\theta}$ برآورد کننده به روش حداکثر درست نمایی برای θ باشد و n به سمت ∞ میل کند. انگاه واریانس $\hat{\theta}$برابر حد پایین نامساوی کرایمر – رائو خواهد شد.

4- MLE دارای خاصیت پایایی است یعنی اگر فرض کنید $\hat{\theta}=(\hat{\theta}_1,\ldots,\hat{\theta}_k)$ که در آن $\hat{\theta}_j=\hat{f}(x_1,\ldots,x_n)$ است که برآوردکننده درستنمایی ماکزیمم $\theta=(\theta_1,\ldots,\theta_4)$ در چگالی $f(\theta_1,\ldots,\theta_n)$ باشد، آنگاه برآوردکننده درستنمایی ماکزیمم $\tau(\theta)=(\tau_1(\theta),\ldots,\tau_r(\theta))$ برابر است با

$$\tau(\hat{\theta})=(\tau_1(\hat{\theta}),\ldots,\tau_r(\hat{\theta}))$$

مثال21: فرض کنید $x\sim N(\mu,1)$ برآورد کننده MLE پارامتر $\alpha=\mu^2$ را پیدا کنید.

$$\hat{\mu}=\bar{x}\Rightarrow\hat{\alpha}=\hat{\mu}^2=\bar{x}^2$$

5- برآوردکننده های کافی

برآورد کننده ای مانند $\hat{\theta}$ را کافی می نامیم در صورتی که از همه اطلاعات یک نمونه، مربوط به برآورد پارامتر θ یک جامعه بهره‌برداری کند. یعنی اگر تمام دانشی را که می توانیم با مشخص کردن مقادیر فردی نمونه ها و ترتیب آنها بدست آوریم، بتوانیم تنها با مشاهده مقدار آماره $\hat{\theta}$ هم بدست آوریم. این را می توان بر حسب توزیع شرطی مقادیر نمونه به فرض $\hat{\theta}$ بیان کرد. این کمیت با عبارت زیر داده می شود.

$$f\left(x_1, x_2......, x_n \middle| \hat{\theta}\right) = \frac{f\left(x_1, x_2...., x_n, \hat{\theta}\right)}{g\left(\hat{\theta}\right)}$$

اگر این عبارت به $\hat{\theta}$ بستگی داشته باشد، مقادیر خاص $x_n, ..., x_2, x_1$ که $\hat{\theta}$ را بدست می دهند برای برخی مقادیر θ محتملتر از سایر مقادیرند و دانستن این مقادیر نمونه ای به برآورد θ کمک خواهد کرد. اگر این عبارت به θ بستگی نداشته باشد، مقادیر خاص $x_n, ..., x_2, x_1$ که $\hat{\theta}$ را بدست می دهند برای هر مقدار θ به یک اندازه محتمل اند و دانستن این مقادیر نمونه ای به برآورد θ کمکی نخواهد کرد.

تعریف: آماره $\hat{\theta}$ یک برآوردکننده کافی پارامتر θ است اگر و تنها اگر به ازای هر مقدار $\hat{\theta}$، توزیع شرطی نمونه تصادفی $x_n, ..., x_2, x_1$ به فرض $\hat{\theta}$، مستقل از θ باشد.

مثال22: اگر $x_n, ..., x_2, x_1$ متغیرهای تصادفی برنولی با پارامتر یکسان θ باشند. نشان دهید که آماره $\hat{\theta} = \frac{X}{n}$ که در آن $X = x_1 + ... + x_n$، برآورد کننده کافی برای θ است.

$$f(x;\theta) = \theta^{x_i}(1-\theta)^{1-x_i} \qquad ; \quad x_i = 0,1$$

$$f(x_1, ..., x_n) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} = \theta^{x}(1-\theta)^{n-x} = \theta^{n\hat{\theta}}(1-\theta)^{n-n\hat{\theta}}$$

X یک متغیر دو جمله ای با پارامترهای n و θ است لذا:

$$b(x;n,\theta)=\binom{n}{x}\theta^{x}(1-\theta)^{n-x}$$

$$g(\hat{\theta})=\binom{n}{n\hat{\theta}}\theta^{n\hat{\theta}}(1-\theta)^{n-n\hat{\theta}} \quad ; \qquad \hat{\theta}=0,\frac{1}{n},......,1$$

$$f(x_1,....,x_n|\hat{\theta})=\frac{f(x_1,....,x_n)}{g(\hat{\theta})}=\frac{\theta^{n\hat{\theta}}(1-\theta)^{n-n\hat{\theta}}}{\binom{n}{n\hat{\theta}}\theta^{n\hat{\theta}}(1-\theta)^{n-n\hat{\theta}}}=\frac{1}{\binom{n}{x}}$$

که بهθ بستگی ندارد.

مثال23: نشان دهید که آماره $y=\frac{1}{6}(x_1+2x_2+3x_3)$ برای برآورد پارامتر θ ی جامعه برنولی، کافی نیست.

باید نشان دهیم $f(x_1,x_2,x_3|y)=\frac{f(x_1,x_2,x_3,y)}{g(y)}$ به ازای برخی مقادیر x_3,x_2,x_1 مستقل از θ نیست. بنابراین حالت خاصی را در نظر می گیریم که در آن $x_3=0,x_2=1,x_1=1$ بطوری که

$$f\left(1,1,0\middle|y=\frac{1}{2}\right)=\frac{P\left(x_1=1,x_2=1,x_3=0,y=\frac{1}{2}\right)}{p\left(y=\frac{1}{2}\right)}=\frac{f(1,1,0)}{f(1,1,0)+f(0,0,1)}=\frac{\theta^2(1-\theta)}{\theta^2(1-\theta)+\theta(1-\theta)^2}=\theta$$

که به θ بستگی دارد.

قضیه: آماره θ یک برآوردکننده کافی پارامتر θ است اگر و فقط اگر چگالی یا توزیع احتمال توام نمونه تصادفی را بتوان تجزیه کرد بطوری که

$$f(x_1,....,x_n;\theta)=g(\hat{\theta},\theta)h(x_1,....,x_n)$$

که در آن $g(\hat{\theta},\theta)$ تنها به $\hat{\theta}$ و θ بستگی دارد و $h(x_1,....,x_n)$ به θ بستگی ندارد.

مثال24: نشان دهید که $\bar{x}$ یک برآورد کننده کافی μ جامعه نرمال یا واریانس معلوم σ^2 است.

$$f(x_1, x_2, \ldots, x_n, \mu) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n . e^{-\frac{1}{2}\sum\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

$$\sum(x_i-\mu)^2 = \sum(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2$$

$$f(x_1, \ldots, x_n; \mu) = \left\{\frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\bar{x}-\mu}{\sigma\sqrt{n}}\right)^2}\right\} \times \left\{\frac{1}{\sqrt{n}}\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{n-1} . e^{-\frac{1}{2}.\sum\left(\frac{x_i-\bar{x}}{\sigma}\right)^2}\right\}$$

که در آن اولین عامل سمت راست تنها به برآورد $\bar{x}$ و به میانگین جامعه، μ، بستگی دارد و دومین عامل شامل μ نیست. بنابراین $\bar{x}$ یک برآورد کننده کافی μ جامعه نرمال با واریانس معلوم σ^2 است.