فصل دهم

S-۱

۱- برآورد

در فصل قبل با نمونه‌گیری از یک جامعه با معلوم بودن تابع چگالی و پارامترهای جامعه مثل میانگین و واریانس آشنا شدید. همچنین نحوه بدست آوردن احتمالات مربوط به آماره‌های خاص مثل میانگین، نسبت واریانسها و ..... را نشان دادیم.

در عمل بسیاری ار اوقات می‌دانیم که یک جامعه مثلاً نرمال می‌باشد اما مقدار دقیق پارامترهای جامعه را که μ و σ^2 می‌باشند نمی‌دانیم، در این حالت با نمونه‌گیری از جامعه و با استنباط آماری روی نمونه‌ها می‌توانیم مقادیر پارامترهای مجهول جامعه را با دقت زیادی بدست بیاوریم. بنابراین در این فصل به بحث پیرامون روشهای برآوردیابی می‌پردازیم.

۲ . ۱ . ۱ برآورد نقطه‌ای

با نمونه‌گیری از جامعه دو پارامتر میانگین نمونه‌ها و واریانس نمونه‌ها را بدست می‌آوریم. این دو پارامتر نقش اصلی را در برآورد میانگین و واریانس جامعه بازی می‌کنند. بطور کلی با استفاده از میانگین و واریانس نمونه‌ها به دو طریق می‌توان پارامترهای مجهول جامعه را برآورد نمود. در روش اول که به برآورد نقطه‌ای معروف است با برابر قرار دادن میانگین و واریانس نمونه‌ها با میانگین جامعه، این دو پارامتر مجهول را بدست می‌آوریم. اما در روش دوم که به برآورد فاصله‌ای معروف است برای پارامتر مجهول جامعه مثل میانگین یک بازه با استفاده از پارامترهای معلوم نمونه بدست می‌آوریم و نشان می‌دهیم که با احتمال زیاد پارامتر مجهول جامعه در این بازه قرار دارد.

در برآورد نقطه‌ای از دو روش زیر استفاده می‌کنیم که در ادامه با آنها آشنا می‌شوید:

۱- برآورد به روش گشتاورها

۲- برآورد به روش حداکثر احتمال (در ...... ماکزیمم).

۳ ۱۰ . ۱ . ۱ برآورد به روش گشتاورها

قبلاً kامین گشتاور متغیر تصادفی X را بصورت زیر معرفی نمودیم:

$$m_k = E\left[X^K\right]$$

به همین ترتیب می‌توانیم kامین گشتاور نمونه‌ها را بصورت زیر معرفی کنیم:

$$m'_k = \frac{1}{n}\sum_{i=1}^{n} X_i{}^k$$

n: تعداد نمونه‌ها

با توجه به اینکه kامین گشتاور متغیر تصادفی X به تمام مقادیر احتمال متغیر X وابسته است و همچنین kامین گشتاور نمونه‌ها نیز به هر یک از مقادیر نمونه‌ها وابسته است بنابراین این طور بنظر می‌رسد که با استفاده از گشتاورها بتوان بنحوی پارامترهای مجهول جامعه را تقریب زد.

در این روش به این صورت عمل می‌کنیم که اگر p عدد پارامتر مجهول داشته باشیم به ترتیب گشتاور اول نمونه را با گشتاور اول متغیر تصادفی x، گشتاور دوم نمونه را با گشتاور دوم متغیر تصادفی و گشتاور pام نمونه را با گشتاور pام متغیر تصافی X برابر قرار می‌دهیم به این ترتیب دستگاه معادلات زیر بدست می‌آید:

$$\begin{cases} m_1 = m'_1 \\ m_2 = m'_2 \\ \quad\vdots \\ m_k = m'_k \end{cases} \qquad k = 1, 2, \ldots, p$$

اگر $\theta_1, \theta_2, \ldots, \theta_p$ مقادیر مجهول پارامترها باشد با حل دستگاه فوق هر یک از مقادیر θ_i بدست می‌آید.

حال به مثال زیر توجه کنید:

S-۲

۴ مثال ۱: می‌دانیم تعداد مراجعه کنندگان به یک پمپ بنزین در طول روز یک متغیر تصادفی پواسون با پارامتر مجهول λ می‌باشد. اگر در طول ۱۰ روز مشاهدات زیر را برای تعداد مراجعات به پمپ بنزین بدست آورده باشیم مطاوبست تعیین پارامتر مجهول λ؟

۱۲۰ , ۹۰ , ۸۵ , ۱۱۰ , ۱۳۵ , ۹۵ , ۱۰۰ , ۱۱۱ , ۱۱۵ , ۹۸

حل: برای متغیر تصادفی پواسون داریم:

$$m_1 = E[X'] = E[X] = \lambda$$

از آنجا که تنها یک پارامتر مجهول λ داریم استفاده از اولین گشتاور کفایت می‌کند. حال اولین گشتاور نمونه را بدست می‌آوریم:

$$m'_1 = \frac{1}{n} \sum_{i=1}^{n} X_i = \overline{X} = \frac{1}{10} (120 + 90 + 85 + 110 + 135 + 95 + 100 + 111 + 115 + 98) = 105/9$$

حال با برابر قرار دادن اولین گشتاور نمونه با اولین گشتاور متغیر تصادفی X مقدار λ بدست می‌آید:

$$m_1 = m'_1 = 105/9 \quad \Rightarrow \quad \lambda = 105/9$$

۵ مثال ۲: در یک سری مسابقات تیراندازی n مرتبه به هدف شلیک می‌ند که احتمال اصابت گلوله به هدف p می‌باشد. نتایج حاصل از ۱۰ مرتبه شرکت تیرانداز در مسابقات بصورت زیر می‌باشد: (نتایج بر حسب تعداد دفعات اصابت گلوله به هدف می‌باشند)

۷ , ۵ , ۶ , ۸ , ۸ , ۶ , ۷ , ۴ , ۹ , ۸

مطلوبست برآورد پارامترهای مجهول n و p؟

ابتدا توجه می‌نیم که در این مثال تعداد پیروزی‌ها در n مرتبه انجام آزمایشهای مستقل برنولی با احتمال پیروزی p مورد نظر است. بنابراین متغیر تصادفی X دو جمله‌ای با پارامترهای n و p می‌باشد.

در این مثال دو پارامتر مجهول داریم بنابراین از گشتاور مرتبه اول و دوم استفاده می‌کنیم:

$$m_1 = E[X] = np$$

$$m_2 = E[X^2] = \sigma^2 + \mu^2 = npq + n^2p^2 = n(n-1)p + np = np(np + 1 - p)$$

توجه کنید که مقادیر اول و دوم گشتاور متغیر تصادفی X همواره با استفاده از میانگین و واریانس متغیر تصادفی X بدست می‌آیند. به عبارتی در برآورد به روش گشتاورها مستقل از اینکه توزیع متغیر تصادفی X چه باشد همواره داریم:

$$\begin{cases} m_1 = \mu \\ m'_1 = \overline{X} \end{cases} \Rightarrow \quad \mu = \overline{X}$$

$$\begin{cases} m_2 = \sigma^2 + \mu^2 \\ m'_2 = \frac{1}{n} \sum_{i=1}^{n} \overline{X}_i^2 \end{cases} \Rightarrow \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (\overline{X}_i - \overline{X})^2 = \frac{1}{n} \sum_{i=1}^{n} \overline{X}_i^2 - \overline{X}^2$$

یعنی برآورد گشتاوری میانگین و واریانس هر توزیعی برابر است با میانگین و واریانس نمونه‌ها. به همین دلیل $\overline{X}$, $\frac{1}{n} \sum (X_i - \overline{X})^2$ برآوردهای توزیع – آزاد نامیده می‌شوند، یعنی برآوردهایی که مستقل از توزیع تصادفی X می‌باشند.

۶ حال مقادیر گشتاور اول و دوم نمونه را بدست می‌آوریم:

$$m'_1 = \frac{1}{n} \sum_{i=1}^{n} X_i = \frac{1}{10} (7+5+6+8+8+6+7+4+9+8) = 6/8$$

$$m'_2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 = \frac{1}{10} (49+25+36+64+64+36+49+16+81+64) = 48/4$$

به این ترتیب دستگاه معادلات زیر بدست می‌آید:

$$\begin{cases} m_1 = m'_1 & \Rightarrow \quad np = ۶/۸ \\ m_۲ = m'_۲ & \Rightarrow \quad np = (np + ۱ - p) = ۴۸/۴ \end{cases}$$

با حل دستگاه بدست می‌آوریم:

$$۶/۸\,(۶/۸ + ۱ + p) = ۴۸/۴ \quad \Rightarrow \quad ۵۳/۰۴ - ۶/۸\, p = ۴۸/۴ \quad \Rightarrow \quad p = ۰/۶۸$$

$$np = ۶/۸ \quad \rightarrow \quad n\,(۰/۶۸) = ۶/۸ \Rightarrow n = ۱۰$$

بنابراین تیرانداز فوق در هر مسابقه می‌بایستی حدوداً ۱۰ مرتبه به هدف شلیک کرده باشد که احتمال موفقیت وی در هر شلیک ۰/۶۸ می‌باشد.

برآورد به روش گشتاورها همواره نتایج دقیق و رضایت بخشی نمی‌دهد به مثال بعد در این زمینه توجه کنید:

۷ مثال ۳: فرض کنید متغیر تصادفی X در بازه $[\circ, a]$ بصورت یکنواخت توزیع شده باشد در این صورت مقدار a را به ازای نمونه‌های بدست آمده زیر برآورد کنید:

الف) ۵ و ۴ و ۳ و ۱ و ۸ و ۵ و ۴ و ۲ و ۰ و ۱

ب) ۵ و ۱۰ و ۴۰

حل: برای متغیر تصادفی پیوسته یکنواخت داریم:

$$m_۱ = E[X] = \int_\circ^a x \, \frac{۱}{a}\, dx = \frac{۱}{a} \left(\frac{x^۲}{۲}\right)\Big|_\circ^a = \frac{a}{۲}$$

همینطور برای نمونه‌ها داریم:

$$m'_۱ = \frac{۱}{n} \sum_{i=۱}^{n} X_i = \frac{۱}{۱۰}(۱+\circ+۲+۴+۵+۸+۱+۳+۴+۵) = ۳/۳$$

بنابراین در حالت کلی:

$$a = ۲\overline{X} \quad \Leftarrow \quad m'_۱ = \overline{X} = \frac{a}{۲}$$

یعنی برآورد a عبارتست از دو برابر میانگین نمونه‌های بدست آمده که در اینجا با توجه به نمونه‌های الف a برابر می‌شود با: $a = ۲\overline{X} = ۲\times۳/۳ = ۶/۶$.

حال توجه کنید که $a = ۶/۶$ یعنی نمونه‌ها در اصل از بازه $[\circ\,, ۶/۶]$ انتخاب شده‌اند در حالی که در بین نمونه‌ها عدد ۸ موجود می‌باشد که داخل بازه فوق نمی‌باشد و این یعنی برآورد به روش گشتاورها دارای کمبودهایی می‌باشد. همین حالت برای نمونه‌های (ب) نیز صادق است:

ب)

$$\overline{X} = \frac{۱}{۳}(۴۰+۱۰+۵) = ۱۸/۳ \quad \Rightarrow \quad a = ۲\overline{x} = ۳۶/۶$$

بنابراین برآورد به روش گشتاورها در بعضی موارد نتایج دلخواه را بدست نمی‌دهد در نتیجه می‌بایستی از معیاری استفاده کنیم که میزان کارایی یک روش برآورد را نشان دهد تا بتوان میان روشهای مختلف برآورد، بهترین روش را برای مسایل مختلف انتخاب نمود.

S-۳

۲.۱.۱. برآورد به روش حداکثر احتمال

متغیر تصادفی X با تابع چگالی احتمال $f_X(x)$ و پارامترهای مجهول $\theta_۱, \theta_۲, \ldots, \theta_p$ را در نظر بگیرید. از این متغیر تعداد n نمونه استخراج می‌کنیم که عبارتند از $x_۱, x_۲, \ldots, x_n$ حال $X_۱, X_۲, \ldots, X_n$ را متغیرهای تصادفی متناظر با نمونه‌ها در نظر می‌گیریم.

می‌دانیم تابع چگالی احتمال توام متغیرهای تصادفی $X_۱, \ldots, X_n$ مقادیر احتمال وقوع نمونه‌ها را بدست می‌دهد. مقادیری از θ_iها که تابع چگالی احتمال توام X_iها را ماکزیمم می‌کند برآورد پارامترهای مجهول θ_i می‌باشد. زیرا به ازای ماکزیمم شدن تابع چگالی احتمال توام X_iها در واقع احتمال وقوع نمونه‌ها به بیشترین مقدار خود می‌رسد. برای بدست آوردن برآورد θ_iها فرض می‌کنیم $L(\theta_۱, \theta_۲, \ldots, \theta_p)$ تابع چگالی احتمال توام متغیرهای تصادفی X_i باشد در این صورت با توجه به اینکه X_iها از یکدیگر مستقل باشند تابع چگالی احتمال توام آنها عبارتست از:

$$L(\theta_۱, \theta_۲, \ldots, \theta_p) = f_{X_۱}(x_۱) \,.\, f_{X_۲}(x_۲) \cdots f_{X_n}(x_n)$$

حال با مشتق‌گیری از $L(\theta_1, \ldots, \theta_p)$ نسبت به θ_iها و برابر صفر قرار دادن این مشتق‌ها دستگاهی از معادلات بدست می‌آید که با حل این دستگاه مقادیر هر یک از θ_iها بدست می‌آید. معمولاً در محاسبات برای سادگی از $Ln(L)$ مشتق می‌گیریم. به مثال زیر توجه کنید:

۹ مثال ۴: نمونه‌های $X_1, X_2, \ldots, X_n$ را از متغیر تصادفی برنولی X داریم مطلوبست: برآورد پارامتر p به روش حداکثر احتمال (MLE)؟

حل: تابع چگالی متغیر تصادفی برنولی X به صورت زیر می‌باشد:

$$f_X(x) = p^x (1-p)^{1-x} \qquad x = 0, 1$$
$$= 0 \qquad \text{سایر مقادیر}$$

به این ترتیب تابع چگالی احتمال توام نمونه‌ها عبارتست از:

$$L(p) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdots f_{X_n}(x_n) = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i}$$

$$p^{\sum_{i=1}^{n} x_i} (1-p)^{\sum_{i=1}^{n}(1-x_i)} \quad \Rightarrow \quad L(p) = p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i}$$

حال برای اینکه محاسبات ساده‌تر شوند از طرفین رابطه بالا Ln می‌گیریم:

$$Ln(L(p)) = Ln\left(p^{\sum x_i} (1-p)^{n-\sum x_i}\right)$$
$$= \sum x_i \, Ln(p) + (n - \sum x_i)\, Ln(1-p)$$

فرض می‌کنیم $H(p) = Ln(L(p))$ در این صورت می‌بایستی معادله $\frac{dH}{dP} = 0$ را حل کنیم تا مقدار مجهول p بدست آید. بنابراین:

$$\frac{dH}{dP} = 0 \quad \Rightarrow \quad \sum x_i \left(\frac{1}{p}\right) - (n - \sum x_i)\left(\frac{1}{1-p}\right) = 0$$

همانطور که ملاحظه می‌کنید با بکار بردن Ln در محاسبات، مشتق‌گیری بسیار ساده‌تر می‌شود. حال داریم:

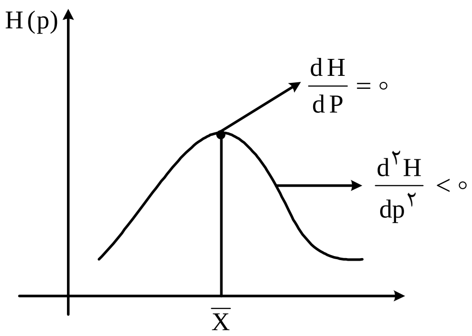
$$\sum x_i \left(\frac{1}{p}\right) = (n - \sum x_i)\left(\frac{1}{1-p}\right)$$
$$\Rightarrow \frac{1-p}{p} = \frac{n - \sum x_i}{\sum x_i} \quad \Rightarrow \quad \frac{1}{p} - 1 = \frac{n}{\sum x_i} - 1$$
$$\Rightarrow \frac{1}{p} = \frac{n}{\sum x_i} \quad \Rightarrow \quad \hat{p} = \frac{1}{n} \sum x_i$$

($\hat{p}$ برآورد پارامتر مجهول p می‌باشد).

بنابراین برآورد پارامتر مجهول p برابر است با $\overline{X} = \frac{1}{n} \sum x_i$ توجه کنید مقدار $\hat{p} = \overline{X}$ بدست آمده تنها در صورتی جواب درستی می‌باشد که نقطه $\overline{X}$ واقعاً مقدار ماکزیمم تابع $H(P)$ باشد. مثلاً ممکن است این نقطه، نقطه مینیمم تابع $H(P)$ باشد. برای اطمینان از اینکه $\hat{p} = \overline{X}$ مقدار ماکزیمم $H(P)$ می‌باشد باید شرط زیر برقرار باشد:

$$\frac{d^2 H}{d p^2} < 0$$

یعنی تقعر تابع $H(P)$ حول $\hat{p} = \overline{X}$ روبه پایین باشد به این ترتیب $\hat{p} = \overline{X}$ نقطه ماکزیمم تابع می‌باشد. در این رابطه به شکل زیر توجه کنید:



حال داریم:

$$\frac{d^2H}{dp^2} = \sum x_i \left(\frac{-1}{p^2}\right) - \left(n-\sum x_i\right)\left(\frac{1}{(1-p)^2}\right)$$

که به ازای هر مقدار p منفی است. بنابراین $\overline{X}$ به عنوان برآورد p در متغیر تصادفی برنولی می‌باشد.

۱۱ حال برآورد p را با استفاده از روش گشتاورها بدست می‌آوریم و نتیجه را با روش حداکثر احتمال مقایسه می‌کنیم:

مطابق روش گشتاورها داریم:

$$m_1 = p$$

$$\Rightarrow \quad m_1 = m'_1 \;\Rightarrow\; p = \overline{X}$$

$$m'_1 = \frac{1}{n}\sum x_i$$

همانطور که ملاحظه می‌کنید نتیجه برآورد از هر دو روش یک نتیجه را بدست می‌دهد. بطور کلی از آنجا که روش حداکثر معمولاً برآوردهای بهتری می‌دهد، هرگاه نتیجه دو روش برآورد با یکدیگر برابر نبود، نتیجه روش برآورد حداکثر احتمال را ملاک قرار می‌دهیم.در برآورد به روش گشتاورها نشان دادیم که این روش برای مواردی مثل مثال ۳ برآورد مطلوبی ارایه نمی‌کند بنابراین در مثال بعدی مثال ۳ را با روش برآورد حداکثر احتمال حل می‌کنیم:

S ۴

۱۲ مثال ۵: متغیر X یک متغیر تصادفی یکنواخت در بازه $(0\,,\,a)$ می‌باشد. مطلوبست برآورد پارامتر مجهول a به روش حداکثر احتمال؟

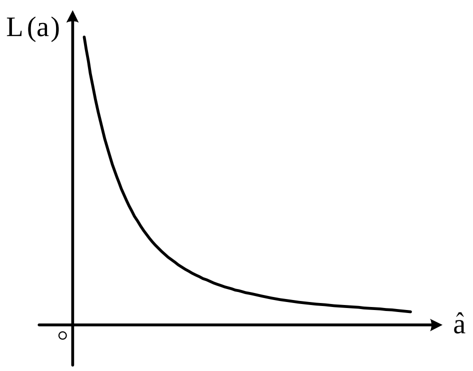
حل: تابع چگالی احتمال X عبارتست از:

$$f_X(x) = \frac{1}{a} \qquad \text{برای} \quad 0 < x < a$$

با در نظر گرفتن یک نمونه تصادفی به اندازه n تابع چگالی احتمال توام بصورت زیر بدست می‌آید:

$$L(a) = \frac{1}{a}\cdot\frac{1}{a}\cdots\frac{1}{a} = \frac{1}{a^n}$$

منحنی نمایش $L(a)$ را در شکل زیر مشاهده می‌کنید.



از روی منحنی پیداست که شیب منحنی در هیچ نقطه‌ای صفر نمی‌شود بنابراین مقدار مشتق $L(a)$ در هیچ نقطه‌ای صفر نمی‌شود. اما با توجه به نمودار می‌بینیم که هر چقدر a به صفر نزدیک می‌شود مقدار $L(a)$ بیشتر می‌شود، و به عبارتی $L(a)$ به ازای کوچکترین مقدار a ماکزیمم می‌شود، از طرفی کوچکترین مقدار a نباید از بزرگترین نمونه بدست آمده کمتر باشد بنابراین برآورد a به روش حداکثر احتمال در این مثال برابر است با بزرگترین نمونه بدست آمده یعنی nامین آماره مرتب $\hat{a} = X_{(n)}$.

بوضوح این برآورد کاراتر از برآورد به روش گشتاورها می‌باشد.

۱۳ مثال ۶: اگر متغیر تصادفی X یک متغیر تصادفی پواسون با پارامتر λ باشد مطلوبست برآورد λ به روش حداکثر احتمال و به روش گشتاورها؟

حل: تابع چگالی متغیر تصادفی پواسون عبارتست از:

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x_i} \qquad x = 0,1,2,\ldots$$

$$= 0 \qquad \text{سایر مقادیر}$$

$$L(\lambda) = \frac{e^{-\lambda}\lambda^{x_1}}{x_1!}\cdot\frac{e^{-\lambda}\lambda^{x_2}}{x_2!}\cdots\frac{e^{-\lambda}\lambda^{x_n}}{x_n!} = \frac{e^{-n\lambda}\lambda^{\sum x_i}}{x_1!\,x_2!\cdots x_n!}$$

$$H(\lambda) = Ln(L(\lambda)) = -n\lambda + \sum x_i\ Ln(\lambda) - Ln(x_1!\,x_2!\cdots x_n!)$$

$$\frac{dH}{d\lambda} = -n + \sum x_i\left(\frac{1}{\lambda}\right) = 0 \quad \Rightarrow \quad \hat{\lambda} = \frac{1}{n}\sum x_i = \overline{X}$$

حال این مساله را به روش گشتاورها حل می‌کنیم:

$$m_1 = \mu = \lambda$$

$$\Rightarrow \quad m_1 = m_1' \ \Rightarrow \ \hat{\lambda} = \overline{X}$$

$$m_1' = \frac{1}{n}\sum x_i = \overline{X}$$

ملاحظه می‌کنید که نتایج هر دو روش در این حالت نیز یکسان می‌باشند.

۱۴ مثال ۷: اگر متغیر تصادفی X یک متغیر تصادفی نرمال با میانگین μ و واریانس σ^2 باشد مطلوبست برآورد پارامترهای μ و σ^2 به روشهای حداکثر احتمال و گشتاورها؟

حل: تابع چگالی احتمال متغیر تصادفی X عبارتست از:

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

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

$$H(\mu,\sigma^2) = Ln\ (L((\mu,\sigma^2)) = Ln\ \left(\frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}}\right) + Ln\ \left(e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}}\right)$$

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

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

$$\frac{\partial H}{\partial \mu} = -\frac{1}{2}(-2) = \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial H}{\partial \sigma^2} = -\frac{n}{2}(\frac{1}{\sigma^2}) - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\,(\frac{-1}{(\sigma^2)^2})$$

$$= \frac{-n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

حال می‌بایستی معادلات زیر را حل کنیم:

$$\begin{cases} \frac{1}{\hat\sigma^2}\sum (x_i-\hat\mu) = 0 \\ -\frac{n}{2\hat\sigma^2}+\frac{1}{2(\hat\sigma^2)^2}\sum (x_i-\mu)^2 = 0 \end{cases}$$

۱۵ جواب معادله اول عبارتست از:

$$\sum (x_i-\hat\mu) = 0 \quad \Rightarrow \quad \sum x_i - \sum \hat\mu = 0 \quad \rightarrow \quad \sum x_i - n\hat\mu = 0$$

$$\Rightarrow \hat\mu = \frac{1}{n}\sum x_i = \overline{X}$$

با قرار دادن $\hat\mu = \overline{X}$ در معادله دوم داریم:

$$-\frac{n}{2\hat\sigma^2} + \frac{1}{2(\hat\sigma^2)^2}\sum (x_i-\overline{X})^2 = 0$$

$$n\,2\hat\sigma^2 + \sum (x_i-\overline{X})^2 = 0 \quad \Rightarrow \quad \hat\sigma^2 = \frac{1}{n}\sum (x_i-\overline{X})^2$$

بنابراین برآورد پارامترهای مجهول μ و σ^2 به روش حداکثر احتمال عبارتند از:

$$\hat\mu = \overline{X} \quad , \quad \hat\sigma^2 = \frac{n-1}{n} S^2$$

۱۶ حال برآورد را به روش گشتاورها بدست می‌آوریم:

$$m_1 = \mu$$

$$m_2 = \sigma^2 + \mu^2$$

$$m'_1 = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X}$$

$$m'_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\begin{cases} m_1 = m'_1 \\ m_2 = m'_2 \end{cases} \quad \Rightarrow \quad \begin{cases} \mu = \overline{X} \\ \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \sigma^2 + \mu^2 \end{cases}$$

از معادله اول داریم $\hat\mu = \overline{X}$ با جاگذاری در معادله دوم بدست می‌آید:

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 = \hat{\sigma}^2 + \overline{X}^2 \qquad \Rightarrow \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{X}^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\,\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^2 = \frac{n-1}{n} S^2$$

در نتیجه برآورد پارامترهای μ و σ^2 به روش گشتاورها نیز همان نتایج برآورد به روش حداکثر را بدست می‌دهد.

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2$$

S-۵

۱۶ مثال ۸: اگر متغیر تصادفی X بصورت یکنواخت در بازه $[-b\,,b]$ توزیع شده باشد مطلوبست برآورد b به روش گشتاورها و حداکثر احتمال؟

حل: میانگین متغیر تصادفی X عبارتست از:

$$E[X] = \int_{-b}^{b} dx = \frac{1}{2b}\left(\frac{x^2}{2}\right)\Big|_{-b}^{b} = 0$$

حال طبق روش گشتاورها می‌بایستی از حل معادله $m_1 = m_1'$ پارامتر b بدست آید اما می‌بینیم که:

$$m_1 = E[X] = 0$$

$$m_1' = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X}$$

$$m_1 = m_1' \ \Rightarrow\ \overline{X} = 0$$

همانطور که ملاحظه می‌کنید $\overline{X} = 0$ کمکی در یافتن پارامتر مجهول به ما نمی‌کند. در واقع این مثال حالت خاصی است که در آن جهت تعین پارامتر مجهول می‌بایستی از گشتاور دوم نمونه و متغیر تصادفی X استفاده کنیم:

$$m_2 = E\left[X^2\right] = \int_{-b}^{b} x^2 \frac{1}{2b}\, dx = \frac{1}{2b}\left(\frac{x^3}{3}\right)\Big|_{-b}^{b}$$

$$= \frac{1}{2b}\left(\frac{b^3}{3} + \frac{b^3}{3}\right) = \frac{b^2}{3}$$

به همین ترتیب داریم:

$$m_2' = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \quad \Rightarrow m_2 = m_2' \quad \Rightarrow \quad \frac{b^2}{3} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\Rightarrow b^2 = 3\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) \ \rightarrow\ \hat{b} = \sqrt{3\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)} = \sqrt{3m_2'}$$

بنابراین برآورد پارامتر مجهول b در این مثال برابر $\sqrt{3m_2'}$ می‌باشد.

۱۷ روش برآورد پارامتر b به روش حداکثر احتمال کاملاً مشابه مثال ۵ می‌باشد به این صورت که:

$$f_X(x) = \frac{1}{2b} \qquad\qquad -b < x < b$$

$$L(b) = \frac{1}{2b}.\frac{1}{2b}\cdots\frac{1}{2b} = \frac{1}{(2b)^n}$$

مجدداً از آنجا که $L(b)$ با نزدیک شدن به صفر ماکزیمم می‌شود بنابراین کوچکترین مقداری که b می‌تواند بخود بگیرد برآورد آن می‌باشد و از آنجا که b نمی‌تواند از بزرگترین مقدار نمونه کمتر باشد بنابراین در این مثال نیز برآورد b عبارتست از $\max|X_i|$ یا همان nامین آماره مرتب $X_{(n)}$.

۲ . ۱ . ۱۸ خواص برآورد کننده‌ها

در بخش روشهای گشتاورها و حداکثر احتمال را برای برآورد پارامترهای مجهول جامعه معرفی نمودیم و نشان دادیم که این دو روش همواره برآورد کننده‌های مشابه بدست نمی‌دهند، بنابراین نیازمند یک معیاری برای سنجش میزان کارایی یک برآوردگر نسبت به برآوردگر دیگری می‌باشیم، در این بخش به معرفی خواص برآورد کننده‌ها و روشهایی برای سنجش میزان کارایی آنها می‌پردازیم.

۱ . ۲ . ۱ . ۱۹ متوسط مربع خطا (MSE)

فرض کنید $\hat{\theta}$ برآورد پارامتر θ باشد در این صورت می‌دانیم که $\left|\hat{\theta}-\theta\right|$ مقدار خطای برآورد می‌باشد. معمولاً برای تخمین میزان خطا از مربع $(\hat{\theta}-\theta)$ استفاده می‌شود و از آنجا که در اینجا n نمونه استخراج می‌شود برای هر بار استخراج این n نمونه مقدار $(\hat{\theta}-\theta)^{۲}$ محاسبه می‌شود و میانگین آن ملاک خواهد بود. به این ترتیب معیار متوسط مربع خطا بصورت زیر بدست می‌آید:

$$MSE\ (\hat{\theta}) = E\left[(\hat{\theta}-\theta)^{۲}\right]$$

توجه کنید که همواره در عمل با توجه به استخراج n نمونه و برآورد پارامتر θ مقداری خطا در برآورد پارامتر θ وجود دارد، بنابراین برای اینکه بهترین برآورد را داشته باشیم می‌بایستی خطای برآورد را حداقل کنیم. بوضوح اگر $\hat{\theta}=\theta$ باشد خطای برآورد صفر می‌باشد و بهترین حالت بدست آمده است.

برای اینکه بتوان خطای برآورد را حداقل نمود ابتدا مقدار متوسط مربع خطا را ساده می‌کنیم:

$$MSE = E\left[(\hat{\theta}-\theta)^{۲}\right]= E\left[(\hat{\theta}^{۲}-۲\theta\hat{\theta}-\theta)^{۲}\right]$$
$$=E\left[\hat{\theta}^{۲}\right]- ۲E\left[\theta\hat{\theta}\right]+ E\left[\theta^{۲}\right]$$

حال یک $E^{۲}\left[\hat{\theta}\right]$ به سمت راست عبارت بالا اضافه و کم می‌کنیم:

$$\Rightarrow MSE = E\left[\hat{\theta}^{۲}\right]- E^{۲}\left[\hat{\theta}\right]+ E^{۲}\left[\hat{\theta}\right]- ۲E\left[\theta\hat{\theta}\right]+ E\left[\theta^{۲}\right]$$

توجه کنید که عبارت A همان واریانس $\hat{\theta}$ می‌باشد. می‌دانیم $\hat{\theta}$ خود یک آماره می‌باشد و از آنجا که آماره‌ها خود یک متغیر تصادفی می‌باشند، بنابراین واریانس $\hat{\theta}$ معنی‌دار می‌باشد.

$$\Rightarrow MSE = var(\hat{\theta}) + (E\left[\hat{\theta}\right]-\theta)^{۲}$$

بنابراین متوسط مربع خطا بصورت فوق خلاصه می‌شود که در آن عبارت $E\left[\hat{\theta}\right]-\theta$ در واقع اندازه اختلاف مرکز توزیع $\hat{\theta}$ از θ را نشان می‌دهد.

۲۰ برای اینکه MSE حداقل شود می‌بایستی مقادیر $(E\left[\hat{\theta}\right]-\theta)^{۲}, var(\hat{\theta})$ حداقل شوند. از طرفی می‌دانیم $(E\left[\hat{\theta}\right]-\theta)^{۲}, var(\hat{\theta})$ هر دو عبارتی همواره مثبت می‌باشند بنابراین MSE در صورتی حداقل می‌شود که این دو عبارت حداقل شوند. حال کمترین مقداری را که این دو عبارت بخود می‌پذیرد را محاسبه می‌کنیم:

$$var(\hat{\theta}) = E\left[\hat{\theta}^{۲}\right]-E^{۲}\left[\hat{\theta}\right]$$

اگر $E^{۲}\left[\hat{\theta}\right]=E\left[\hat{\theta}^{۲}\right]$ واریانس $\hat{\theta}$ برابر صفر می‌شود که کمترین مقداری است که یک عبارت مثبت بخود می‌پذیرد. بنابراین $\min(var(\hat{\theta})) = ۰$

همینطور حداقل $(E\left[\hat{\theta}\right]-\theta)^{۲}$ با توجه به مثبت بودن عبارت زمانی رخ می‌دهد که کل عبارت صفر شود یعنی:

$$\min\left[(E\left[\hat{\theta}\right]-\theta)^{۲}\right]= ۰ \quad \Rightarrow \quad (E\left[\hat{\theta}\right]-\theta)^{۲} = ۰ \quad \Rightarrow \quad E\left[\hat{\theta}\right]= \theta$$

می‌دانیم در عمل $var(\hat{\theta})$ هرگز صفر نمی‌شود (مگر اینکه تمام مقادیر نمونه‌ها برابر باشند که آن هم بی‌معنی است). اما رخ دادن تساوی $E\left[\hat{\theta}\right] = \theta$ امکان‌پذیر است بنابراین از میان برآوردگرهای مختلف برای θ برآوردگری مناسب‌تر است که شرایط زیر را داشته باشد.

۱- واریانس برآوردگر $(\hat{\theta})$ کمترین مقدار ممکن را داشته باشد.

۲- $E\left[\hat{\theta}\right] = \theta$ باشد.

شرط دوم به برآورد کننده ناریب معروف است که به صورت زیر تعریف می‌شود:

برآورد کننده نااریب: برآورد کننده $\hat{\theta}$ برای θ نااریب نامیده می‌شود اگر $E\left[\hat{\theta}\right]=\theta$.

برابر تساوی $E\left[\hat{\theta}\right]=\theta$ به این معنی است که اگر مثلاً kبار از جامعه تعداد n نمونه استخراج کنیم و به ازای هر بار استخراج n نمونه مقدار $\hat{\theta}$ را محاسبه کنیم و در نهایت میانگین $\hat{\theta}_i$ها ($E\left[\hat{\theta}\right]=\frac{1}{k}\sum_{i=1}^{k}\hat{\theta}_i$) می‌بایستی برابر با پارامتر مجهول θ شود. برای روشن شدن مطلب به مثال زیر توجه کنید.

21 **مثال ۹:** آیا برآوردگر بدست آمده برای توزیع نرمال، برنولی، پواسون نااریب می‌باشد؟

حل: قبلاً نشان دادیم که برآوردگر μ و σ^2 در توزیع نرمال با استفاده از روشهای گشتاورها و حداکثر احتمال عبارتست از:

$$\hat{\mu}=\overline{X}$$

$$\hat{\sigma}^2=\frac{n-1}{n}S^2$$

برای $\hat{\mu}$ داریم:

$$E[\hat{\mu}]=E[\overline{X}]=\mu$$

توجه کنید که قبلاً نشان دادیم برای متغیر تصادفی X با هر توزیعی $E[\overline{X}]$ برابر μ می‌باشد.

بنابراین $\hat{\mu}=\overline{X}$ در توزیع نرمال یک برآورد نااریب برای μ می‌باشد.

حال برای $\hat{\sigma}^2$ داریم:

$$E\left[\hat{\sigma}^2\right]=E\left[\frac{n-1}{n}S^2\right]=\frac{n-1}{n}E\left[S^2\right]$$

22 در فصل قبل نشان دادیم که اگر $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})$ آنگاه داریم:

$$E\left[S^2\right]=\sigma^2$$

بنابراین:

$$E\left[\hat{\sigma}^2\right]=\frac{n-1}{n}\sigma^2$$

ملاحظه می‌کنید که برای $\hat{\sigma}^2$ مقدار $E\left[\hat{\sigma}^2\right]$ برابر با σ^2 نمی‌باشد بنابراین برآوردگر $\hat{\sigma}^2=\frac{n-1}{n}S^2$ برآوردگر نااریبی برای σ^2 نمی‌باشد. در این حالت می‌گوییم برآوردگر $\hat{\sigma}^2$ یک برآوردگر اریب برای σ^2 می‌باشد.

توجه کنید که $\hat{\sigma}^2$ تنها به دلیل وجود ضریب $(\frac{n-1}{n})$ اریب می‌باشد، به عبارتی تعداد نمونه‌ها عامل اصلی اریب بودن این برآوردگر می‌باشد بنابراین اگر تعداد نمونه‌ها زیاد باشد $(n\to\infty)$ داریم:

$$\lim_{n\to\infty}(\frac{n-1}{n})\sigma^2=\sigma^2$$

یعنی با شرط $n\to\infty$ برآوردگر $\hat{\sigma}^2$ یک برآوردگر نااریب برای σ^2 می‌باشد. به این نوع برآوردگرها، برآوردگر نااریب مجانبی می‌گویند که بصورت زیر تعریف می‌شود:

برآوردگر نااریب مجانبی: اگر $\hat{\theta}$ یک برآوردگر اریب برای θ بر اساس نمونه تصادفی nتایی باشد، می‌گوییم $\hat{\theta}$ یک برآوردگر نااریب مجانبی برای θ است هرگاه داشته باشیم:

$$\lim_{n\to\infty}E\left[\hat{\theta}\right]=\theta$$

23 برای توزیع برنولی قبلاً بدست آوردیم که:

$$\hat{p}=\overline{X} \quad\Rightarrow\quad E[\hat{p}]=E[\overline{X}]=\mu=p$$

همینطور برای توزیع پواسون داریم:

$$\hat{\lambda}=\overline{X}$$

$$E\left[\hat{\lambda}\right]=E[\overline{X}]=\mu=\lambda$$

یک نتیجه کلی که از برآوردگر نااریب بدست می‌آوریم این است که هرگاه پارامتر مجهول توزیع برابر با میانگین توزیع متغیر تصادفی X باشد

(یعنی $\theta = E[X] = \mu$) همینطور از آنجا که برای هر توزیعی مستقل از نوع توزیع داریم $E[\overline{X}] = \mu$ در این صورت $\overline{X}$ همواره به عنوان یک برآوردگر ناالریب این گونه از توزیعها، منظور می‌شود.

خاصیت ناالریبی اگر چه یک معیار برای محک بهتر بودن یک برآوردگر می‌باشد اما از آنجا که برای یک پارامتر می‌توان چندین برآوردگر ناالریب بدست آورد، بنابراین می‌بایستی معیارهای دیگری وجود داشته باشند تا بتوان با کمک آنها از میان برآوردگرهای ناالریب بهترین را انتخاب نمود. برای این منظور معیار کارایی یا موثرتر بودن را معرفی می‌کنیم:

تعریف: اگر $\hat{\theta}_1, \hat{\theta}_2$ هر دو برآوردگرهایی ناالریب برای پارامتر θ باشند در این صورت می‌گوییم $\hat{\theta}_1$ کاراتر یا موثرتر از $\hat{\theta}_2$ می‌باشد هرگاه:

$$\mathrm{var}(\hat{\theta}_1) < \mathrm{var}(\hat{\theta}_2)$$

علت استفاده از این معیار با توجه به تعریف متوسط مربع خطا (MSE) کاملاً قابل توجیه است.

۲۴ مثال ۱۰: از یک جامعه نرمال با میانگین μ و واریانس σ^2 یک نمونه به حجم n انتخاب شده است. در این صورت کدامیک از برآوردگرهای ناالریب زیر معیار مناسب‌تری برای برآورد پارامتر μ می‌باشد؟

الف) $\hat{\mu}_1 = \overline{X}$

ب) $\hat{\mu}_2 = \frac{1}{3}(X_1 + X_3 + X_5)$

حل: با توجه به اینکه هر دو برآوردگر ناالریب می‌باشند بنابراین واریانس آنها را محاسبه نموده و با یکدیگر مقایسه می‌کنیم:

$$\mathrm{var}(\hat{\mu}_1) = \mathrm{var}(\overline{X}) = \frac{\sigma^2}{n}$$

$$\mathrm{var}(\hat{\mu}_2) = \mathrm{var}\left(\frac{1}{3}(X_1 + X_3 + X_5)\right) = \frac{1}{9}(\mathrm{var}(X_1) + \mathrm{var}(X_3) + \mathrm{var}(X_5))$$

$$= \frac{1}{9}(3\sigma^2) = \frac{\sigma^2}{3}$$

با مقایسه دو واریانس بدست آمده داریم:

$$\frac{\sigma^2}{n} < \frac{\sigma^2}{3} \qquad \text{برای } n > 3$$

بنابراین اگر تعداد نمونه‌های استخراج شده از متغیر تصادفی X بیستر از ۳ باشد خواهیم داشت:

$$\mathrm{var}(\hat{\mu}_1) < \mathrm{var}(\hat{\mu}_2)$$

و در نتیجه با توجه به معیار کارایی می‌توان نتیجه گرفت که برآورد کننده $\overline{X}$ بهتر از $\frac{1}{3}(X_1 + X_3 + X_5)$ در این مثال می‌باشد.

تعریف: برای دو برآوردگر $\hat{\theta}_1, \hat{\theta}_2$ نسبت $\frac{\mathrm{var}(\hat{\theta}_1)}{\mathrm{var}(\hat{\theta}_2)}$ را کارایی نسبی برآوردگر $\hat{\theta}_2$ نسبت به برآوردگر $\hat{\theta}_1$ می‌نامند.

بوضوح اگر مقدار کسر از یک بیشتر باشد می‌گوییم برآوردگر $\hat{\theta}_2$ از برآوردگر $\hat{\theta}_1$ کاراتر یا موثرتر است.

S-۷

۲۵ مثال ۱۱: از یک جامعه نمونه‌ای به حجم ۲۰ با میانگین $\overline{X}_1 = 15$ و واریانس $\sigma_1^2 = 4$ استخراج کرده‌ایم. همچنین نمونه دیگری به حجم ۳ با میانگین $\overline{X}_2 = 180$ و واریانس $\sigma_2^2 = 2$ استخراج کرده‌ایم. اگر $\overline{X}$ را به عنوان برآوردگر میانگین جامعه انتخاب کنیم کارایی نسبی $\overline{X}_1$ را محاسبه کنید؟

حل: ابتدا واریانس $\overline{X}_1$ و $\overline{X}_2$ را محاسبه می‌کنیم:

$n_۱ = ۲۰\ ,\quad \overline{X}_۱ = ۱۵\quad ,\quad \sigma_۱^۲ = ۴$

$n_۲ = ۳۰\ ,\quad \overline{X}_۲ = ۱۸\quad ,\quad \sigma_۱^۲ = ۲$

$$\operatorname{var}(\overline{X}_۱) = \frac{\sigma_۱^۲}{n_۱} = \frac{۴}{۲۰} = \frac{۱}{۵}$$

$$\operatorname{var}(\overline{X}_۲) = \frac{\sigma_۲^۲}{n_۲} = \frac{۲}{۳۰} = \frac{۱}{۱۵}$$

$$\overline{X}_۲ \text{ نسبت به } \overline{X}_۱ \text{ کارایی نسبی } = \frac{\operatorname{var}(\overline{X}_۲)}{\operatorname{var}(\overline{X}_۱)} = \frac{\frac{۱}{۱۵}}{\frac{۱}{۵}} = \frac{۵}{۱۵} = \frac{۱}{۳}$$

بنابراین استفاده از نمونه متوسط استخراج شده دوم برای برآورد میانگین جامعه کارایی بیشتری دارد و موثرتر می‌باشد. همچنین توجه کنید که در محاسبه کارایی از متوسط دو نمونه استخراج شده استفاده‌ای نمی‌شود.

۲۶ تا بحال نشان دادیم که یک برآوردگر به شرطی قابل قبول است که در درجه اول ناااریب باشد و همچنین کمترین واریانس را در میان سایر برآوردگرها داشته باشد بنابراین تعریف زیر را داریم:

تعریف: اگر $\hat{\theta}$ یک برآورد کننده نااریب برای θ باشد. در این صورت $\hat{\theta}$ بهترین برآورد کننده نااریب خطی برای θ می‌باشد. هرگاه:

۱- $\hat{\theta}$ تابعی خطی از $X_۱, X_۲, \dots, X_n$ باشد.

۲- برای هر θ $\quad E\left[\hat{\theta}\right] = \theta$.

۳- از میان تمام برآورد کننده‌های $\hat{\theta}_i$ که در شرایط ۱ و ۲ صدق می‌کند داشته باشیم:

$$\operatorname{var}(\hat{\theta}) \leq \operatorname{var}(\hat{\theta}_i)$$

با توجه به استفاده از روش متوسط مربع خطا (MSE) تعریف بالا و شرایط آن کاملاً موجه می‌باشد. همچنین توجه کنید که اگر $\hat{\theta}$ در شرایط فوق صدق کند در این صورت $\hat{\theta}$ موثرترین برآورد کننده نااریب خطی نیز می‌باشد.

حال به مثال زیر توجه کنید:

۲۷ مثال ۱۲: فرض کنید X یک متغیر تصادفی با میانگین μ و واریانس $\sigma^۲$ باشد. هرگاه $X_۱, X_۲, \dots, X_n$ یک نمونه تصادفی از X باشد نشان دهید توزیع متغیر تصادفی X هر چه باشد، برآورد کننده $\overline{X}$ بهترین برآورد کننده نااریب خطی برای μ می‌باشد؟

حل: ابتدا فرض می‌کنیم $\hat{\theta}$ یک برآورد کننده نااریب خطی برای μ باشد در این صورت $\hat{\theta}$ می‌بایستی به فرم زیر باشد:

$$\hat{\theta} = \sum_{i=۱}^{n} (a_i\ x_i) + b$$

حال متغیرهای a_i و b را طوری بدست می‌آوریم که $\hat{\theta}$ بهترین برآورد کننده برای μ باشد. برای این منظور مطابق تعریف $\hat{\theta}$ می‌بایستی در شرایط تعریف بهترین برآورد کننده نااریب خطی صدق کند. یعنی: $E\left[\hat{\theta}\right] = \mu$ بنابراین:

$$E\left[\hat{\theta}\right]=E\left[\sum_{i=۱}^{n}(a_i\ x_i)+b\right]=E\left[\sum_{i=۱}^{n}a_i\ x_i\right]+b$$

$$=E\sum_{i=۱}^{n}(a_i\,E[X_i])+b=\sum_{i=۱}^{n}(a_i\ m_۱)+b$$

$$=m_۱\sum_{i=۱}^{n}a_i+b=E[X]\sum_{i=۱}^{n}a_i+b\qquad\Rightarrow\ E\left[\hat{\theta}\right]=\mu\sum_{i=۱}^{n}a_i+b$$

حال چون بایستی داشته باشیم $E\left[\hat{\theta}\right]=\nu$ بنابراین:

$$\mu\sum_{i=۱}^{n}a_i+b=\mu\quad\Rightarrow\quad\begin{cases}\sum_{i=۱}^{n}a_i=۱\\ b=۰\end{cases}$$

۲۸ حال می‌بایستی واریانس برآورد کننده $\hat{\theta}$ را بدست آورده و سپس آنرا حداقل کنیم:

$$\mathrm{var}(\hat{\theta})=\mathrm{var}(\sum_{i=۱}^{n}a_i X_i+b)=\mathrm{var}(\sum a_i X_i)$$

$$=\sum\mathrm{var}(a_i X_i)=\sum a_i^۲\ \mathrm{var}(X_i)=\sum a_i^۲\ \sigma^۲=\sigma^۲\sum_{i=۱}^{n}a_i^۲$$

بنابراین برای حداقل نمودن واریانس $\sum a_i^۲$ می‌بایستی حداقل شود به عبارتی باید مقادیر $a_۱,a_۲,\ldots,a_n$ را طوری پیدا نمود که در دو شرط زیر صدق کنند:

۱- $\sum_{i=۱}^{n}a_i=۱$

۲- $\sum a_i^۲$ مینیمم شود.

برای حداقل نمودن $\sum a_i^۲$ ابتدا مقدار $a_۱$ را از شرط اول محاسبه نموده و حاصل را در شرط دوم $(\sum a_i^۲)$ قرار می‌دهیم. به صورت زیر:

$$\sum_{i=۱}^{n}a_i=۱\qquad\Rightarrow\qquad a_۱+\sum_{i=۲}^{n}a_i=۱\qquad\Rightarrow\qquad a_۱=۱-\sum_{i=۲}^{n}a_i$$

فرض می‌کنیم $Q=\sum_{i=۱}^{n}a_i^۲$ در این صورت باید Q مینیمم شود پس:

$$Q=\sum_{i=۱}^{n}a_i^۲=a_۱^۲+\sum_{i=۲}^{n}a_i=(۱-\sum_{i=۲}^{n}a_i)^۲+\sum_{i=۲}^{n}a_i^۲$$

$$\Rightarrow Q=(۱-\sum_{i=۲}^{n}a_i)^۲+\sum_{i=۲}^{n}a_i^۲$$

حال برای حداقل نمودن Q می‌بایستی مشتقات پاره‌ای Q را نسبت به a_iها محاسبه نموده و برابر صفر قرار دهیم تا مقادیری بشکل $a_۱^*,a_۲^*,\ldots,a_n^*$ که Q را مینیمم می‌کنند بدست آید:

$$\frac{\partial Q}{\partial a_j}=۲(-۱)(۱-\sum_{i=۱}^{n}a_i)+۲a_j\qquad j=۲,۳,\cdots,n$$

۲۹ توجه کنید که مقدار a_1^* از روی سایر a_i^*ها $(i=2,3,\cdots,n)$ بشکل زیر بدست می‌آید:

$$a_1^* = 1-\sum_{i=2}^{n} a_i^*$$

رابطه ۱-۱۰

$$\frac{\partial Q}{\partial a_j} = 0 \quad \Rightarrow \quad \begin{cases} a_j^* = (1-\sum\limits_{i=2}^{n} a_i^*) = a_1^* \\ a_1^* = 1-\sum\limits_{i=2}^{n} a_i^* \end{cases}$$

حال با حل دستگاه فوق مقادیر a_j^*ها بدست می‌آید، برای حل ابتدا کل معادلات را جداگانه نوشته و طرفین معادلات را جمع می‌کنیم. به این ترتیب داریم:

$$a_1^* = 1-\sum_{i=2}^{n} a_i^*$$

$$a_2^* = 1-\sum_{i=2}^{n} a_i^*$$

$$a_3^* = 1-\sum_{i=2}^{n} a_i^*$$

$$\vdots$$

$$a_n^* = 1-\sum_{i=2}^{n} a_i^*$$

$$\text{حاصل جمع معادلات} = \sum_{i=1}^{n} a_i^* = n - n\sum_{i=2}^{n} a_i^* = n\,\underbrace{(1-\sum_{i=2}^{n} a_i^*)}_{a_i^*} = n a_1^*$$

بنابراین بدست می‌آید:

$$\begin{cases} \sum\limits_{i=1}^{n} a_i^* = n a_1^* \\ \sum\limits_{i=1}^{n} a_i^* = 1 \end{cases} \quad \Rightarrow \quad n a_1^* = 1 \;\Rightarrow\; a_1^* = \frac{1}{n}$$

و با توجه به رابطه ۱-۱۰ بدست می‌آید:

$$(j=1,2,\cdots,n) \quad a_j^* = \frac{1}{n}$$

بنابراین $\sum\limits_{i=1}^{n} a_i^2$ و با شرط $\sum\limits_{i=1}^{n} a_i = 1$ به ازای $a_j^* = \frac{1}{n}$ حداقل می‌شود. بنابراین برآوردگر $\hat{\theta}$ برابر می‌شود با: $\hat{\theta} = \sum\limits_{i=1}^{n} \frac{1}{n} X_i = \frac{1}{n}\sum\limits_{i=1}^{n} X_i = \overline{X}$

در نتیجه $\overline{X}$ بهترین برآوردگر ناریب خطی برای μ می‌باشد و این مساله مستقل از توزیع متغیر تصادفی X می‌باشد. به همین دلیل این مطلب از اهمیت فراوانی برخوردار است از طرفی می‌توان نشان داد که اگر برآوردگر $\hat{\theta}$ تابعی خطی از X_iها هم نباشد باز هم $\overline{X}$ دارای کوچکترین واریانس می‌باشد. حال می‌توان گفت که $\overline{X}$ برای پارامتر μ در تابع نرمال و برای پارامتر p در توزیع برنولی و برای پارامتر λ در توزیع پواسون بهترین برآوردگر ناریب می‌باشد.

۳۰ . ۱۰ . ۳ برآورد کننده سازگار

یکی دیگر از خواصی که برای برآورد کننده‌ها می‌توان بدست آورد خاصیتی است که به سازگاری معروف است. اگر داشته باشیم:

$$\lim_{n\to\infty} MSE(\hat{\theta}) = \lim E\left[(\hat{\theta}-\theta)^{۲}\right] = ۰$$

یعنی به ازای تعداد زیادی نمونه از متغیر تصادفی X مقدار متوسط مربع خطا برابر صفر شود، از طرفی طبق نامساوی مارکف (رجوع کنید به قضیه چبی شف) داریم:

$$p(g(x) \geq k) \leq \frac{E[g(x)]}{k} \Rightarrow ۱-p(g(x)\leq k) \leq \frac{E[g(x)]}{k}$$

$$\Rightarrow p(g(x) \geq k) \geq ۱-\frac{E[g(x)]}{k}$$

حال اگر $k=\varepsilon^{۲}$ $(\varepsilon > ۰)$, $g(x)=(\hat{\theta}-\theta)^{۲}$ باشد در این صورت نامساوی بصورت زیر خواهد بود:

$$p((\hat{\theta}-\theta)^{۲} \leq \varepsilon^{۲}) \geq ۱-\frac{E\left[(\hat{\theta}-\theta)^{۲}\right]}{\varepsilon^{۲}}$$

$$\Rightarrow p(|\hat{\theta}-\theta| < \varepsilon) \geq ۱-\frac{E\left[(\hat{\theta}-\theta)^{۲}\right]}{\varepsilon^{۲}}$$

حال اگر شرط $\lim_{n\to\infty} MSE = ۰$ برقرار باشد می‌بایستی $E\left[(\hat{\theta}-\theta)^{۲}\right]$ برابر صفر باشد در این صورت نامساوی بصورت زیر خلاصه می‌شود:

$$p(\left|\hat{\theta}-\theta\right| < \varepsilon) \geq ۱-\frac{۰}{\varepsilon^{۲}} \Rightarrow p(\left|\hat{\theta}-\theta\right| < \varepsilon) \geq ۱$$

و از آنجا که احتمال حداکثر برابر ۱ واحد می‌باشد بنابراین:

$$p(\left|\hat{\theta}-\theta\right| < \varepsilon) = ۱ \qquad \forall \varepsilon > ۰$$

یا به بیان دیگر:

$$p(\left|\hat{\theta}-\theta\right| \geq \varepsilon) = ۰ \qquad \forall \varepsilon > ۰$$

در صورتی که احتمالات فوق برای یک برآورد کننده $\hat{\theta}$ برقرار باشند می‌گوییم برآورد کننده دارای خاصیت سازگاری می‌باشد.

بوضوح خاصیت سازگاری زمانی که تعداد نمونه‌ها به اندازه کافی بزرگ باشد نشان دهنده برتری یک برآورد کننده نسبت به دیگری می‌باشد. اما در حالتی که تعداد نمونه‌ها کم باشد، سازگار بودن یک برآورد کننده معیاری برای بهتر بودن آن نمی‌باشد.

قبل از ارایه یک مثال به قضیه زیر نیز توجه کنید:

۳۱ قضیه: اگر داشته باشیم: $\lim_{n\to\infty} E\left[\hat{\theta}\right] = \theta$ و $\lim_{n\to\infty} var\left[\hat{\theta}\right] = ۰$ آنگاه $\hat{\theta}$ یک برآورد کننده سازگار برای θ می‌باشد. این قضیه از آنجا نتیجه می‌شود که در تعریف سازگاری داشتیم:

$$\lim_{n\to\infty} MSE = ۰ \Rightarrow \lim_{n\to\infty} E[(\hat{\theta}-\theta)^{۲}] = ۰$$

$$\Rightarrow \lim_{n\to\infty} (var(\hat{\theta}) + (E[\hat{\theta}]-\theta)^{۲}) = ۰ \Rightarrow \begin{cases} \lim_{n\to\infty} var(\hat{\theta}) = ۰ \\ \lim_{n\to\infty} E[\hat{\theta}] = \theta \end{cases}$$

۳۲ مثال ۱۳: اگر X یک متغیر تصادفی با میانگین μ و واریانس $\sigma^{۲}$ باشد نشان دهید توزیع X هر چه باشد $\overline{X}$ یک برآورد کننده سازگار از روی نمونه‌های $X_{۱}, X_{۲}, \ldots, X_{n}$ برای μ می‌باشد؟

حل: می‌دانیم:

$$E[\overline{X}] = \mu$$

$$var(\overline{X}) = \frac{\sigma^۲}{n}$$

به این ترتیب:

$$\lim_{n\to\infty} E[X] = \lim_{n\to\infty} \mu = \mu$$

$$\lim_{n\to\infty} var(\overline{X}) = \lim_{n\to\infty} \frac{\sigma^۲}{n} = ۰$$

در نتیجه $\overline{X}$ یک برآورد کننده سازگار برای μ می‌باشد.

۳۳ مثال: اگر نمونه‌های $X_۱, X_۲, \ldots, X_n$ را از یک جامعه نرمال با میانگین μ و واریانس $\sigma^۲$ استخراج کنیم در این صورت نشان دهید که $S^۲$ یک برآورد کننده سازگار برای $\sigma^۲$ می‌باشد؟

حل: قبلاً نشان دادیم که $\frac{(n-۱)S^۲}{\sigma^۲}$ یک متغیر تصادفی خی دو با $n-۱$ درجه آزادی است $(\chi^۲_{n-۱})$ و بنابراین میانگین آن $n-۱$ و واریانس آن $۲(n-۱)$ می‌باشد:

$$E\left[\frac{(n-۱)S^۲}{\sigma^۲}\right] = n-۱$$

$$var(\frac{(n-۱)S^۲}{\sigma^۲}) = ۲(n-۱)$$

حال با ساده نمودن روابط فوق داریم:

$$\frac{n-۱}{\sigma^۲} E[S^۲] = n-۱ \quad \Rightarrow \quad E[S^۲] = \frac{n-۱}{n-۱}\sigma^۲ \quad \Rightarrow \quad E[S^۲] = \sigma^۲$$

$$var(\frac{(n-۱)\, S^۲}{\sigma^۲}) = ۲(n-۱) \quad \Rightarrow \quad \frac{(n-۱)^۲}{\sigma^۴} var(S^۲) = ۲(n-۱)$$

$$var(S^۲) = \frac{۲\sigma^۴}{n-۱} \quad \Rightarrow \quad \lim_{n\to\infty} var(S^۲) = \lim_{n\to\infty} \frac{۲\sigma^۴}{n-۱} = ۰$$

بنابراین $S^۲$ یک برآورد کننده سازگار برای $\sigma^۲$ می‌باشد.

S-۹

۳۴ ۱۰ . ۴ آماره کافی

اگر بتوان یک آماره یافت که تمام اطلاعات را در بازه یک پارامتر مجهول خلاصه کند آنگاه آنرا یک آماره کافی برای پارامتر مجهول می‌نامیم. آماره‌های کافی به صورت زیر تعریف می‌شود:

تعریف: اگر نمونه تصادفی $X_۱, X_۲, \ldots, X_n$ را از یک متغیر تصادفی با پارامتر مجهول θ داشته باشیم در این صورت آمارهٔ $y=g(X_۱,X_۲,\ldots,X_n)$ را یک آماره کافی برای θ می‌نامیم اگر و فقط اگر برای هر مقدار $Y=y$ توزیع مشروط هر آماره بفرم $h(X_۱,X_۲,\ldots,X_n)$ و با شرط $Y=y$ به θ وابسته نباشد.

به این معنی که اگر مقدار آماره کافی معلوم باشد $(Y=y)$ آنگاه به خود مقادیر نمونه نیازی نیست و مقادیر نمونه اطلاعات بیشتری در بازه پارامتر مجهول θ نسبت به آماره کافی بدست آمده، نمی‌دهند.

۳۵ مثال ۱۴: فرض می‌کنیم $X_۲,X_۱$ یک نمونه از متغیر تصادفی برنولی با پارامتر مجهول p باشند در این صورت نشان دهید که $Y=X_۱+X_۲$ یک آماره کافی برای p می‌باشد؟

حل: تابع احتمال متغیر تصادفی برنولی بفرم زیر است:

$$p_X(x) = p^x (۱-p)^{۱-x} \qquad x=۰,۱$$
$$= ۰ \qquad \text{سایر مقادیر}$$

حال تعریف می‌کنیم:

$Y=X_۱+X_۲=g(X_۱,X_۲)$ باید نشان دهیم اگر $V=h(X_۱,X_۲)$ یک آماری دیگر باشد در این صورت $P_{V,Y}(V|Y=y)$ به θ وابسته نمی‌باشد. برای این منظور ابتدا لازم است توزیع توام V و Y را محاسبه کنیم که عبارتست از:

$$P_{V,Y}(V=h(۰,۰)\ ,\ Y=۰) = (۱-p)\,(۱-p) = (۱-p)^۲$$
$$P_{V,Y}(V=h(۰,۱)\ ,\ Y=۱) = p\,(۱-p)$$
$$P_{V,Y}(V=h(۱,۰)\ ,\ Y=۰) = (۱-p)\,p$$
$$P_{V,Y}(V=h(۱,۱)\ ,\ Y=۲) = p.p = p^۲$$

حال توجه کنید که توزیع $Y=X_۱+X_۲$ یک توزیع دو جمله‌ای با پارامتر $n=۲$ و p می‌باشد بنابراین: $P_Y(y) = \binom{۲}{y} p^y (۱-p)^{۲-y}$.

مقدار توزیع مشروط برای هر یک از مقادیر $Y=y$ عبارتست از:

$$P_{V,Y}(V=h(۰,۰)\,|\,Y=۰) = \frac{(۱-p)^۲}{\binom{۲}{۰} p^۰ (۱-p)^{۲-۰}} = ۱$$

$$P_{V,Y}(V=h(۰,۱) \text{ یا } h(۱,۰)\,|\,Y=۱) = \frac{p(۱-p)}{\binom{۲}{۱} p(۱-p)} = \frac{۱}{۲}$$

$$P_{V,Y}(V=h(۱,۱)\,|\,Y=۲) = \frac{p^۲}{\binom{۲}{۲} p^۲ (۱-p)^۰} = ۱$$

بنابراین اگر V هر آماره دلخواهی باشد، تابع احتمال مشروط آن به p وابسته نمی‌باشد در نتیجه $Y=X_۱+X_۲$ یک آماره کافی برای پارامتر مجهول p می‌باشد.

۳۶ محاسبه کافی بودن یک آماره با توجه به تعریف کار بسیار مشکلی است، اما با استفاده از روش تجزیه فیشر- نیمن یک راه نسبتاً ساده برای اثبات کافی بودن یک آماره وجود دارد:

قضیه: $\hat{\theta}$ را یک برآورگر کافی θ می‌نامیم اگر و فقط اگر تابع چگالی احتمال توام نونه‌ها را بتوان بفرم زیر تجزیه کرد:

$$f(x_۱,x_۲,\dots,x_n.\theta) = \prod_{i=۱}^{n} f_X(x_i) = g(\hat{\theta},\theta)\,.\,h(x_۱,x_۲,\dots,x_n)$$

مثال ۱۵: اگر یک نمونه nتایی از یک متغیر تصادفی برنولی داشته باشیم نشان دهید $Y=\sum_{i=۱}^{n} X_i$ یک آماره کافی برای p می‌باشد؟

حل:

$$\prod_{i=1}^{n} P_X(x_i) = \prod_{i=1}^{n} \left[p^{x_i} (1-p)^{1-x_i}\right] = p^{\sum x_i} (1-p)^{n-\sum x_i}$$

حال داریم:

$$\begin{cases} g(\hat{\theta},\theta) = p^{\sum x_i} (1-p)^{1-x_i} \\ h(x_1, x_2, \ldots, x_n) = 1 \end{cases}$$

بنابراین $Y = \sum_{i=1}^{n} X_i$ یک آماره کافی برای p می‌باشد.

۳۷ توجه کنید هرگاه $\hat{\theta}$ یک برآورد کننده نااریب برای θ بر حسب یک آماره کافی باشد آنگاه می‌توان نشان داد که $\hat{\theta}$ دارای کمترین واریانس در میان سایر برآورد کننده‌های نااریب برای θ می‌باشد. به همین دلیل است که برآورد کننده‌ها را بر پایه آماره‌های کافی بدست می‌آورند.

قبلاً نشان دادیم که $(\frac{n-1}{n})\, S^2$ یک برآورد کننده نااریب برای σ^2 متغیر تصادفی نرمال می‌باشد، حال با استفاده از برآورد کننده کافی یک برآورد کننده نااریب و کافی برای متغیر تصادفی نرمال بدست می‌آوریم.

برای متغیر تصادفی نرمال با میانگین μ و واریانس σ^2 داریم:

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

$$= g(\hat{\theta},\theta)\,.\, h(x_1, x_2, \cdots, x_n)$$

که در آن:

$$h(x_1, x_2, \cdots, x_n) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}}$$

از طرفی به جای μ نیز می‌توان از برآورد کافی آن $\overline{X}$ استفاده نمود. در نتیجه $\sum (x_i - \overline{X})^2$ یک آماره کافی برای σ^2 می‌باشد. از طرفی قبلاً نشان دادیم که:

$$E\left[\frac{1}{n}\sum (x_i - \overline{X})^2\right] = \frac{n-1}{n}\sigma^2$$

$$\Rightarrow E\left[\frac{n-1}{n}\underbrace{\left(\frac{1}{n-1}\sum (x_i - \overline{X})^2\right)}_{S^2}\right] = E\left[\frac{n-1}{n} S^2\right] = \frac{n-1}{n}\sigma^2$$

$$\Rightarrow \frac{n-1}{n} E[S^2] = \frac{n-1}{n}\sigma^2 \quad \Rightarrow \quad E[S^2] = \sigma^2$$

بنابراین خود S^2 یک برآورد کننده نااریب و کافی برای σ^2 می‌باشد. و دارای کوچکترین واریانس در میان چنین برآورد کننده‌هایی می‌باشد.

۳۸ . ۱۰ . ۵ خاصیت پایایی برآورد کننده‌ها

یکی از نتایج کلی که از برآورد کننده‌های حداکثر احتمال بدست می‌آید خاصیت پایایی برآورد کننده حداکثر احتمال است. اگر $\hat{\theta}$ برآورد کننده حداکثر احتمال θ باشد و بخواهیم تابعی از θ مثل $\gamma = h(\theta)$ را برآورد کنیم کافیست به جای θ مقدار برآورد شده آنرا که $\hat{\theta}$ می‌باشد جایگزین کنیم در این صورت برآورد γ عبارتست از: $\hat{\gamma} = h(\hat{\theta})$

این مطلب نشان می‌دهد که برای محاسبه برآورد تابعی از پارامتر مجهول نیازی به محاسبه مجدد برآورد کننده نمی‌باشد بلکه کافیست یکبار برآورد برای پارامتر θ بدست آید. توجه کنید که این مطلب فقط برای برآورد کننده‌هایی که به روش حداکثر احتمال بدست می‌آیند صادق می‌باشد.

مثال ۱۶: اگر X یک متغیر تصادفی هندسی نوع اول باشد مطلوبست:

الف) برآورد پارامتر p به روش حداکثر احتمال.

ب) برآورد تابع $h(p) = p(X \le ۲)$ با استفاده از خاصیت پایایی.

حل: تابع احتمال توزیع هندسی نوع اول عبارتست از:

$$P_X(x) = p(۱-p)^x \qquad x = ۰,۱,۲,\cdots$$
$$= ۰ \qquad \text{سایر مقادیر}$$

$$L(p) = \prod_{i=۱}^{n} p(۱-p)^{x_i} = p^n (۱-p)^{\sum x_i}$$

$$H(p) = Ln(L(p)) = n\,Ln(p) + \sum x_i\, Ln(۱-p)$$

$$\frac{dH}{dP} = \frac{n}{p} - \frac{\sum x_i}{۱-\hat{p}} \quad \Rightarrow \quad \frac{۱-\hat{p}}{\hat{p}} = \frac{۱}{n}\sum x_i = \overline{X} \quad \Rightarrow \quad \frac{۱}{\hat{p}} - ۱ = \overline{X} \quad \Rightarrow \quad \hat{p} = \frac{۱}{۱+\overline{X}}$$

بنابراین برآورد p به روش حداکثر احتمال برابر $\frac{۱}{۱+\overline{X}}$ می‌باشد.

ب)

$$p(X \le ۲) = p(X=۰) + p(X=۱) + p(X=۲)$$
$$= p + p(۱-p) + p(۱-p)^۲$$

حال داریم:

$$\gamma = h(p) = p(۱+(۱-p)+(۱-p)^۲)$$

بنابراین خاصیت پایایی $\hat{\gamma} = h(\hat{p})$ بنابراین:

$$\hat{\gamma} = h(\hat{p}) = \frac{۱}{۱+\overline{X}}\left(۱+\left(۱-\frac{۱}{۱+\overline{X}}\right) + \left(۱-\frac{۱}{۱+\overline{X}}\right)^۲\right)$$

S-۱۰

۴۰ ۱۰ . ۶ برآورد فاصله‌ای

در برآورد نقطه‌ای با توجه به نمونه‌های بدست آمده نشان دادیم برای پارامترهای مجهول جامعه می‌توان عددی بدست آورد که با خطای کمی بیانگر مقدار واقعی پارامتر مجهول باشد. اما واقعاً نمی‌دانیم با چه مقدار خطایی پارامتر مجهول را برآورد کرده‌ایم برآورد فاصله‌ای روشی است که در آن برای پارامتر مجهول یک بازه در نظر می‌گیریم و نشان می‌دهیم که با احتمالی معین پارامتر مجهول در بازه مورد نظر قرار دارد.

فرض کنید نمونه‌های $X_۱, X_۲, \ldots, X_n$ را از یک متغیر تصادفی X بدست آورده باشیم در این صورت فاصله $(L_۱, L_۲)$ را طوری می‌سازیم که با احتمال بسیار زیاد مثلاً ۹۵٪ شامل پارامتر مجهول جامعه (مثلاً μ) باشد. به فاصله $(L_۱, L_۲)$ یک فاصله اطمینان می‌گویند زیرا مطمئن هستیم که به احتمال زیاد (مثلاً ۹۵٪) این فاصله شامل پارامتر مجهول می‌باشد. مقادیر $L_۱, L_۲$ هر کدام آماره‌ای هستند که بر اساس مقادیر معلوم نمونه‌ها بدست می‌آیند. حال به تعریف دقیق فاصله اطمینان توجه کنید:

هرگاه مقادیر $X_۱, X_۲, \ldots, X_n$ تعداد n نمونه تصادفی بدست آمده از یک متغیر تصادفی X با پارامتر مجهول θ باشند. در این صورت می‌گوییم دو آماره $L_۱, L_۲$ یک $(۱-\alpha)$ ۱۰۰ درصد فاصله اطمینان برای θ تشکیل می‌دهند. هرگاه:

$$p(L_۱ \le \theta \le L_۲) \ge ۱-\alpha$$

توجه کنید که $۱-\alpha$ مقدار احتمال می‌باشد و بنابراین در بازه ۰ تا ۱ قرار دارد، در نتیجه برای بیان آن بصورت درصد آنرا درصد ضرب کرده‌ایم و به صورت $(۱-\alpha)$ ۱۰۰٪ می‌نویسیم.

در برآورد فاصله‌ای از آماره‌های بدست آمده از فصل قبل استفاده می‌کنیم. اینکه کدام پارامتر جامعه مجهول باشد، تعین می‌کند که از کدام آماری در ساختن فاصله اطمینان استفاده کنیم. مثال بعد روش ساختن فاصله اطمینان برای جامعه نرمال را نشان می‌دهد.

۴۱ مثال ۱۷: در شهر A بارش باران یک متغیر تصادفی نرمال با میانگین μ سانتی‌متر و واریانس ۴ می‌باشد اگر میزان بارش را در طول ۱۶ روز اندازه بگیریم و میانگین ۱۰ سانتی‌متر را بدست بیاوریم، در این صورت با احتمال ۹۵٪ میانگین بارش باران در شهر A در چه فاصله اطمینانی قرار می‌گیرد؟

حل: معلومات مساله عبارتند از:

$n = ۱۶$

$\overline{X} = ۱۰$

$\mu =$ مجهول

$\sigma^{۲} = ۴ \rightarrow \sigma = ۲$

ابتدا توجه کنید که با توجه به جدول متغیر تصادفی نرمال استاندارد رابطه زیر برقرار می‌باشد:

$$p(-۱/۹۶ \le Z \le ۱/۹۶) = ۰/۹۵ \quad \text{(رابطه ۱۰-۲)}$$

زیرا داریم:

$$N_Z(۱/۹۶) = p(Z \le ۱/۹۶) = ۰/۹۷۵$$

$$N_Z(-۱/۹۶) = ۱ - N_Z(۱/۹۶) = ۰/۰۲۵$$

$$\Rightarrow p(-۱/۹۶ \le Z \le ۱/۹۶) = p(Z \le ۱/۹۶) - p(Z \le ۱/۹۶) = ۰/۹۷۵ - ۰/۰۲۵ = ۰/۹۵$$

۴۲ در فصل نمونه‌گیری نشان دادیم که اگر واریانس جامعه معلوم باشد آمارهٔ $Z = \dfrac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$ یک متغیر تصادفی نرمال استاندارد می‌باشد و از طرفی با توجه به رابطه ۱۰ − ۲ چون Z نیز یک متغیر تصادفی نرمال استاندارد می‌باشد می‌توان نوشت:

$$P(-۱/۹۶ \le Z \le ۱/۹۶) = P\left(-۱/۹۶ \le \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \le ۱/۹۶\right)$$

با ساده نمودن نامساوی $-۱/۹۶ \le \dfrac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}} \le ۱/۹۶$ بدست می‌آوریم:

$$\overline{X} - ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \overline{X} + ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right)$$

بنابراین رابطه ۱۰ − ۲ بصورت زیر بدست می‌آید:

$$p\left(\overline{X} - ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \overline{X} + ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right)\right) = ۰/۹۵$$

همانطور که ملاحظه می‌کنید برای μ یک فاصله اطمینان ۹۵ درصدی به شکل $(L_۱, L_۲)$ بدست آوردیم که در آن $L_۱, L_۲$ عبارتند از:

$$L_۱ = \overline{X} - ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$L_۲ = \overline{X} + ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right)$$

حال با جایگذاری $\overline{X}$ و σ و n در رابطه بدست می‌آوریم:

$$L_۱ = ۱۰ - ۱/۹۶\left(\frac{۲}{\sqrt{۱۶}}\right) = ۹/۰۲$$

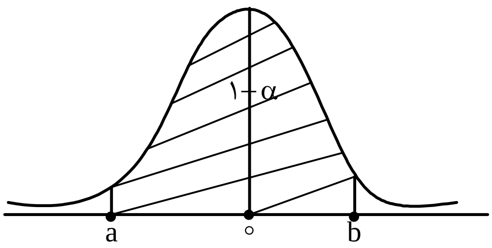
$$L_۲ = ۱۰ + ۱/۹۶\left(\frac{۲}{\sqrt{۱۶}}\right) = ۱۰/۹۸$$

در نتیجه:

$$p(۹/۰۲ \le \mu\ ۱۰۹۸) = ۰/۹۵$$

یعنی ۹۵ درصد اطمینان داریم که میانگین جامعه در بازه $(۱۰/۹۸ , ۹/۰۲)$ قرار دارد حال برای اینکه بتوانیم یک فاطله اطمینان $(۱-\alpha)۱۰۰$ درصدی در حالت کلی برای جامعه نرمال با میانگین مجهول و واریانس معلوم بسازیم ابتدا مفهوم $Z_{۱-\frac{\alpha}{۲}}$ را معرفی می‌کنیم.

۴۳ نمودار زیر منحنی نمایش توزیع نرمال استاندارد می‌باشد ملاحظه می‌کنید که مساحت مشخص شده در شکل برابر $۱-\alpha$ می‌باشد:



برای پیدا کردن فاصله اطمینان می‌بایستی اعداد a و b را طوری پیدا کرد که:

$$p(a \leq Z \leq b) = p\left(a \leq \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq b\right) = ۱-\alpha$$

با فرض اینکه فاصله متقارن باشد مساحت زیر منحنی بعد از نقطه b و قبل از نقطه a برابر $\frac{\alpha}{۲}$ می‌باشد. در نتیجه مساحت کل زیر منحنی قبل از نقطه b برابر است با $۱-\frac{\alpha}{۲} = ۱-\alpha + \frac{\alpha}{۲}$ که آنرا با $Z_{۱-\frac{\alpha}{۲}}$ نمایش می‌دهیم. به این ترتیب $Z_{۱-\frac{\alpha}{۲}}$ نقطه‌ای روی محور افقی است که مساحت زیر منحنی قبل از آن برابر $۱-\frac{\alpha}{۲}$ می‌باشد. این مطلب با توجه به تعریف تابع توزیع نرمال استاندارد بصورت زیر نیز قابل بیان است:

$$p(Z \leq Z_{۱-\frac{\alpha}{۲}}) = ۱-\frac{\alpha}{۲}$$

حال می‌توان به جای b و a در $p(a \leq Z \leq b)$ مقادیر $Z_{۱-\frac{\alpha}{۲}}$ و $-Z_{۱-\frac{\alpha}{۲}}$ را جایگزین نمود و با ساده نمودن نامساوی خواهیم داشت:

$$p(-Z_{۱-\frac{\alpha}{۲}} \leq Z \leq Z_{۱-\frac{\alpha}{۲}}) = p\left(-Z_{۱-\frac{\alpha}{۲}} \leq \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq Z_{۱-\frac{\alpha}{۲}}\right)$$

$$p=\left(-Z_{۱-\frac{\alpha}{۲}}\left(\frac{\sigma}{\sqrt{n}}\right) \leq \overline{X}-\mu \leq Z_{۱-\frac{\alpha}{۲}}\left(\frac{\sigma}{\sqrt{n}}\right)\right)$$

$$p=\left(\overline{X}-Z_{۱-\frac{\alpha}{۲}}\left(\frac{\sigma}{\sqrt{n}}\right) \leq \mu \leq \overline{X}+Z_{۱-\frac{\alpha}{۲}}\left(\frac{\sigma}{\sqrt{n}}\right)\right)=۱-\alpha$$

بنابراین با توجه به آخرین رابطه یک فاصله $(۱-\alpha)۱۰۰$ درصد برای میانگین (μ) جامعه نرمال با معلوم بودن $\sigma^۲$ بدست آوردیم.

۴۴ مثال ۱۸: میزان فروش ماهانه یک فروشگاه یک متغیر تصادفی نرمال با میانگین مجهول μ و انحراف معیار ۵۰ هزار تومان در ماه می‌باشد. نتایج حاصل از ۲۵ ماه فروش نشان می‌دهند که میانگین فروش در طول این ۲۵ ماه عبارتست از ۵۵۰ هزار تومان. اگر ۹۰ درصد اطمینان داشته باشیم میانگین (μ) در فاصله (a,b) می‌باشد مقادیر a و b را مشخص کنید؟

حل: معلومات مثال عبارتند از:

$n = ۲۵$

$\overline{X} = ۵۵۰ \qquad \mu =$ مجهول $\qquad \sigma = ۵۰$

می‌دانیم فاصله اطمینان از رابطه زیر بدست می‌آید:

$$p\,(\overline{X} - Z_{۱-\frac{\alpha}{۲}}\,(\frac{\sigma}{\sqrt{n}}) \le \mu \le \overline{X} + Z_{۱-\frac{\alpha}{۲}}\,(\frac{\sigma}{\sqrt{n}})\,) = ۱-\alpha$$

ابتدا می‌بایستی مقدار α را بدست بیاوریم و سپس از روی آن $Z_{۱-\frac{\alpha}{۲}}$ بدست می‌آید.

$$\%۱۰۰\,(۱-\alpha)\ \%۹۰ \quad \Rightarrow \quad ۱-\alpha = ۰/۹ \quad \Rightarrow \quad \alpha = ۰/۱$$

$$Z_{۱-\frac{\alpha}{۲}} \overset{\alpha = ۰/۱}{=} Z_{۱-\frac{۰/۱}{۲}} = Z_{۰/۹۵}$$

۴۵ با توجه به جدول توزیع نرمال استاندارد $Z_{۰/۹۵}$ عبارتست از:

$$N_Z\,(Z_{۰/۹۵}) = ۰/۹۵ \quad \Rightarrow \quad Z_{۰/۹۵} = ۱/۶۴۵$$

بنابراین بدست می‌آوریم:

$$p(۵۵۰ - ۱/۶۴۵\,(\frac{۵۰}{\sqrt{۲۵}}) \le \mu \le ۵۵۰ + ۱/۶۴۵\,(\frac{۵۰}{\sqrt{۲۵}}) = ۰/۹$$

$$\Rightarrow p\,(۵۳۳/۵۵ \le \mu \le ۵۶۶/۴۵\,) = ۰/۹$$

بنابراین با احتمال ۰/۹ متوسط درآمد فروشگاه در بازه $(۵۶۶/۴۵\ ,\ ۵۳۳/۵۵)$ قرار دارد. توجه کنید که طول بازه فاصله اطمینان در حالت کلی برابر است با:

$$L_۲ - L_۱ = (\overline{X} + Z_{۱-\frac{\alpha}{۲}}\,\frac{\sigma}{\sqrt{n}}) - (\overline{X} - Z_{۱-\frac{\alpha}{۲}}\,(\frac{\sigma}{\sqrt{n}})) \quad \Rightarrow L_۲ - L_۱ = ۲Z_{۱-\frac{\alpha}{۲}}\,(\frac{\sigma}{\sqrt{n}})$$

ملاحظه می‌کنید که طول بازه مستقل از $\overline{X}$ می‌باشد و تنها به تعداد نمونه‌ها وابسته می‌باشد. یعنی می‌توان تعداد نمونه‌ها را طوری انتخاب نمود که هر فاصله اطمینان دلخواهی برای μ بدست آید.

به عنوان مثال در مثال قبل با ۲۵ نمونه طول بازه عبارتست از:

$$L_۲ - L_۱ = ۲(۱/۶۴۵)\,(\frac{۵۰}{\sqrt{۱۰۰}}) = ۱۶/۴۵$$

پیداست که با ۱۰۰ نمونه طول بازه نصف شده است.

۱۰ . ۶ . ۱ فاصله اطمینان برای جامعه نرمال با میانگین مجهول μ و واریانس مجهول $\sigma^۲$. حال فاصله اطمینان را برای حالتی بدست می‌آوریم که علاوه بر میانگین جامعه، واریانس نیز مجهول باشد. از فصل نمونه‌گیری می‌دانیم متغیر تصادفی $T = \dfrac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}}$ یک متغیر تصادفی t با $n-۱$ درجه آزادی می‌باشد بنابراین مجدداً مشابه روشی که قبلاً برای بدست آوردن فاصله اطمینان ارایه کردیم داریم:

$$p(-t_{۱-\frac{\alpha}{۲}} \le t_{(n-۱)} \le t_{۱-\frac{\alpha}{۲}}) = p(-t_{۱-\frac{\alpha}{۲}} \le \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \le t_{۱-\frac{\alpha}{۲}})$$

$$= p\,(\overline{X} - t_{۱-\frac{\alpha}{۲}}\,(\frac{S}{\sqrt{n}}) \le \mu \le \overline{X} + t_{۱-\frac{\alpha}{۲}}\,(\frac{S}{\sqrt{n}})\,) = ۱-\alpha$$

که در آن $t_{1-\frac{\alpha}{2}}$ عبارتست از: $t(n-1)\ (T \le t_{1-\frac{\alpha}{2}}) = 1-\frac{\alpha}{2}$.

بنابراین فاصله اطمینان در این حالت عبارتست از:

$$L_1 = \overline{X} - t_{1-\frac{\alpha}{2}} (\frac{S}{\sqrt{n}})$$

$$L_2 = \overline{X} + t_{1-\frac{\alpha}{2}} (\frac{S}{\sqrt{n}})$$

47 **مثال ۱۹:**میزان مصرف بنزین در هر ماه یک متغیر تصادفی نرمال با کیانگین مجهول μ و واریانس σ^2 می‌باشد اگر اطلاعات زیر از یک نمونه در طول ۶ ماه بدست آمده باشد مطلوبست یک فاصله اطمینان ۹۰ درصدی برای متوسط میزان مصرف بنزین؟

$$\sum_{i=1}^{6} x_i = 246 \qquad \sum_{i=1}^{6} x_i^2 = 11200$$

حل:

$$\overline{X} = \frac{1}{n}\sum_{1}^{n} x_i = \frac{1}{6}(146) = 41$$

$$S^2 = \frac{1}{n-1}(\sum_{1}^{n} (x_i - \overline{X})^2) = \frac{n}{n-1}(\frac{1}{n}\sum_{1}^{n} (x_i - \overline{X})^2)$$

$$= \frac{n}{n-1}(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{X}^2) = \frac{1}{n-1}(\sum_{i=1}^{n} x_i^2 - n\overline{X}^2)$$

$$\Rightarrow S^2 = \frac{1}{5}(11200 - 6(41)^2) = \frac{1}{5}(1114) = 222/8$$

$$\Rightarrow S = 14/92$$

حال برای یک فاصله اطمینان ۹۰ درصدی داریم:

$$1-\alpha = 0/9 \Rightarrow \alpha = 0/1$$

$$t_{1-\frac{\alpha}{2}} = t_{0/95}$$

48 با توجه به جدول t با ۵ درجه آزادی مقدار $t_{0/95}$ برابر ۲/۰۱۵ بدست می‌آید.

در نتیجه:

$$L_1 = \overline{X} - t_{0/95} (\frac{S}{\sqrt{n}}) = 41 - 2/015 (\frac{14/92}{\sqrt{6}}) = 28/72$$

$$L_2 = \overline{X} + t_{0/95} (\frac{S}{\sqrt{n}}) = 41 + 2/015 (\frac{14/92}{\sqrt{6}}) = 53/27$$

بنابراین با ۹۰ درصد اطمینان می‌توانیم بگوییم که متوسط مصرف بنزین ماهانه مابین ۲۸/۷۲ و ۵۳/۲۷ میلیون لیتر خواهد بود.

برای حالتی که σ^2 مجهول باشد طول بازهٔ فاصله اطمینان عبارتست از:

$$L_2 - L_1 = 2t_{1-\frac{\alpha}{2}} (\frac{S}{\sqrt{n}})$$

اما توجه می‌کنید که در این حالت طول بازه به متغیر S نیز وابسته می‌باشد و در نتیجه به هیچ عنوان نمی‌توان اندازه نمونه را طوری انتخاب نمود که دقیقاً یک بازه با طول معین بدست آید.

نکته: قبلاً نشان دادیم که اگر تعداد نمونه‌ها زیاد باشد $(n > ۳۰)$ $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ و $\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ هر دو به توزیع نرمال استاندارد میل می‌کنند، در نتیجه تمام مطالب گفته شده در حالتی که $n > ۳۰$ باشد برای متغیر تصادفی X با هر توزیعی برقرار می‌باشد.

۴۹ ۱۰ . ۶ . ۲ فاصله اطمینان برای واریانس جامعه نرمال با میانگین مجهول μ و واریانس $\sigma^۲$.

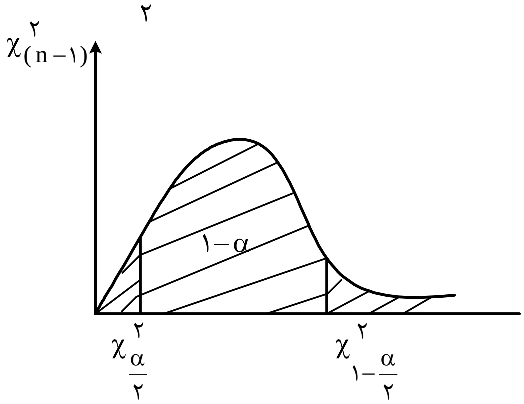
روش بدست آوردن فاصله اطمینان برای واریانس یک جامعه نرمال با میانگین مجهول μ و واریانس $\sigma^۲$ کاملاً مشابه روشی است که برای میانگین جامعه استفاده شد. در این حالت از آماره $\frac{(n-۱)\,S^۲}{\sigma^۲}$ استفاده می‌کنیم. می‌دانیم که $\frac{(n-۱)\,S^۲}{\sigma^۲}$ یک متغیر تصادفی خی دو با $n-۱$ درجه آزادی است.

می‌دانیم:

$$p\,(\chi^۲_{\frac{\alpha}{۲}} \leq \chi^۲_{(n-۱)} \leq \chi^۲_{۱-\frac{\alpha}{۲}}) = ۱-\alpha$$

$$\Rightarrow p\,(\chi^۲_{\frac{\alpha}{۲}} \leq \frac{(n-۱)S^۲}{\sigma^۲} \leq \chi^۲_{۱-\frac{\alpha}{۲}}) = ۱-\alpha \qquad (۳-۱۰)$$

توجه کنید که مقدار ابتدایی بازه $(L_۱)$ برابر $\chi^۲_{\frac{\alpha}{۲}}$ می‌باشد زیرا مطابق شکل منحنی توزیع $\chi^۲_{(n-۱)}$ مساحت زیر منحنی قبل از $\chi^۲_{\frac{\alpha}{۲}}$ برابر $\frac{\alpha}{۲}$ می‌باشد.



به عبارتی $p\,(\chi^۲_{(n-۱)} \leq \chi^۲_{\frac{\alpha}{۲}}) = \frac{\alpha}{۲}$.

حال با ساده نمودن رابطه (۱۰-۳) داریم:

$$p(\frac{(n-۱)\,S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}} \leq \sigma^۲ \leq \frac{(n-۱)\,S^۲}{\chi^۲_{\frac{\alpha}{۲}}}) = ۱-\alpha$$

بنابراین بازه زیر یک فاصله اطمینان $(۱-\alpha)\,۱۰۰$ درصدی برای واریانس بدست می‌دهد:

$$L_۱ = \frac{(n-۱)\,S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}} \qquad , \qquad L_۲ = \frac{(n-۱)\,S^۲}{\chi^۲_{\frac{\alpha}{۲}}}$$

۵۰ مثال ۲۰: کارخانه‌ای ظروف پلاستیکی تولید می‌کند، ضایعات تولید شده توسط ماشین‌های تولیدی یک متغیر تصادفی نرمال با میانگین μ و واریانس $\sigma^۲$ می‌باشد، اگر پراکندگی ضایعات تولید شده در روز زیاد شود (واریانس تعداد ضایعات زیاد شود). می‌بایستی ماشین‌های تولیدی را تعمیر

نمود. برای تخمین واریانس یک نمونه ۲۵ تایی از تعداد ضایعات تولید شده در هر روز می‌گیریم و نتایج $\sum x_i = ۱۷۰$ ، $\sum x_i^۲ = ۲۷۵۰$ را بدست می‌آوریم.

در صورتی که یک فاصله اطمینان ۹۰ درصدی برای واریانس بدست بیاوریم و طول فاصله اطمینان بیشتر از ۵۰ شود می‌بایستی ماشینهای تولیدی را تعمیر کنیم، آیا نیازی به تعمیر می‌باشد؟

حل: معلومات مثال عبارتند از:

$$n = ۲۵$$

$$\overline{X} = \frac{۱}{۲۵}(۱۷۰) = ۶/۸$$

$$S^۲ = \frac{۱}{n-۱}\left(\sum x_i{}^۲ - n\overline{X}^۲\right) = \frac{۱}{۲۴}\left(۲۷۵۰ - ۲۵(۶/۸)^۲\right) = ۶۶/۴۱$$

$$\Rightarrow S = ۸/۱۴$$

$$۱-\alpha = ۰/۹ \quad \rightarrow \quad \alpha = ۰/۱ \quad \rightarrow \quad \chi^۲_{\frac{\alpha}{۲}}(n-۱) = \chi^۲_{۰/۰۵}(۲۴) = ۱۳/۸$$

$$\chi^۲_{۱-\frac{\alpha}{۲}}(n-۱) = \chi^۲_{۰/۹۵}(۲۴) = ۳۶/۴$$

بنابراین:

$$L_۱ = \frac{(n-۱)\,S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}} = \frac{(۲۴)\,(۶۶/۴۱)}{۳۶/۴} = ۴۳/۷۸$$

$$L_۲ = \frac{(n-۱)\,S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}} = \frac{۲۴\,(۶۶/۴۱)}{۱۳/۸} = ۱۱۵/۴۹$$

$$L_۲ - L_۱ = ۱۱۵/۴۹ - ۴۳/۷۸ = ۷۱/۷$$

با توجه به اینکه طول بازه فاصله اطمینان بیشتر از ۵۰ واحد شده است بنابراین می‌بایستی ماشینهای تولیدی تعمیر شوند.

۵۱ . ۱۰ . ۶ . ۳ فاصله اطمینان برای متغیر تصادفی نمایی با پارامتر مجهول

می‌دانیم متغیر تصادفی نمایی عموماً مدلی برای طول عمر وسایل الکتریکی، تجهیزات و ... می‌باشد. بنابراین اگر بخواهیم اطمینان داشته باشیم که متوسط طول عمر یک متغیر تصادفی نمایی حداقل برابر تعداد ساعات معین است در این صورت می‌بایستی $\frac{۱}{\lambda}$ از عدد معین بیشتر باشد بنابراین برای متغیر تصادفی نمایی تنها یک فاصله اطمینان یک طرفه می‌سازیم.

هرگاه $X_۱, X_۲, \ldots, X_n$ یک نمونه تصادفی از یک متغیر تصادفی نمایی با پارامتر λ باشد آنگاه:

$$p\left(\lambda \le \frac{\chi^۲_{۱-\alpha}}{۲n\overline{X}}\right) = ۱-\alpha$$

یک فاصله اطمینان $(۱-\alpha)$ ۱۰۰ درصدی برای λ تشکیل می‌دهد که در آن $\chi^۲_{۱-\alpha}$ مقداری از توزیع $\chi^۲$ با $۲n$ درجه آزادی است که مساحت زیر منحنی نمایش آن برابر $۱-\alpha$ باشد.

۵۲ مثال ۲۱: طول عمر یک مدار الکتریکی یک متغیر تصادفی نمایی با پارامتر λ می‌باشد تعداد ۲۰ عدد از این مدارها را آزمایش می‌کنیم تا زمانی که از کار بیفتند. و ملاحظه می‌کنیم که مجموع طول آنها برابر ۲۳۶۰ ساعت می‌شود مطلوبست: تخمین متوسط طول عمر مدار مربوط با فاصله اطمینان ۹۰ درصدی:

حل: معلومات مساله عبارتند از:

$$n = ۲۰$$

$$\sum_{i=۱}^{۲۰} x_i = ۲۳۶۰ \rightarrow \overline{X} = \frac{۱}{n} \sum_{۱}^{n} x_i = \frac{۱}{۲۰} (۲۳۶۰) = ۱۱۸$$

$$۱-\alpha = ۰/۹ \quad \rightarrow \quad \chi^{۲}_{۱-\alpha} (۲n) = \chi^{۲}_{۰/۹} (۴۰) = ۵۱/۸$$

بنابراین یک فاصله اطمینان یک طرفه ۹۰ درصدی برای λ عبارتست از:

$$p\ (\lambda \le \frac{\chi^{۲}_{۱-\alpha}}{۲n\overline{X}}) = ۱-\alpha \ \Rightarrow\ L_{۲} = \frac{۵۱/۸}{۲(۲۰)\ (۱۱۸)} = ۰/۰۱۰۹$$

توجه کنید که مقدار متوسط طول عمر برای توزیع نمایی $E[X] = \frac{۱}{\lambda}$ می‌باشد بنابراین حداقل متوسط طول عمر مدار عبارتست از:

$$\frac{۱}{۰/۰۱۰۹} = ۹۱/۱۱$$

به عبارتی ۹۰ درصد اطمینان داریم که هر مدار حداقل ۹۱/۱۱ ساعت کار می‌کند.

S-۱۳

۵۳ ۱۰ . ۶ . ۴ فاصله اطمینان برای نسبت p با تعداد نمونه‌های زیاد در توزیع برنولی

نمونه‌های $X_۱, X_۲, \ldots, X_n$ را از یک متغیر تصادفی برنولی با پارامتر مجهول p استخراج می‌کنیم اگر تعداد نمونه‌ها زیاد باشد در این صورت می‌دانیم متغیر تصادفی $\overline{X}$ تقریباً بصورت نرمال با میانگین p و واریانس $\frac{p(۱-p)}{n}$ توزیع می‌شود. به این ترتیب می‌توان نشان داد که یک فاصله اطمینان $(۱-\alpha)۱۰۰$ درصدی برای p عبارتست از:

$$p\ (\ \overline{X} - Z_{۱-\frac{\alpha}{۲}} \sqrt{\frac{\overline{X}(۱-\overline{X})}{n}} \le p \le \overline{X} + Z_{۱-\frac{\alpha}{۲}} \sqrt{\frac{\overline{X}(۱-\overline{X})}{n}}\ ) = ۱ - \alpha$$

توجه کنید که در محاسبه مقدار $\overline{X}$ برابر است با نسبت تعداد موفقیتها به کل تعداد نمونه‌ها به عنوان مثال اگر از ۲۰ مرتبه تیراندازی به هدف ۱۵ بار هدف مورد اصابت قرار گیرد در این صورت $\overline{X} = \hat{p} = \frac{۱۵}{۲۰}$ خواهد بود.

۵۴ مثال ۲۲: احتمال اینکه یک بازیکن فوتبال یک ضربه پنالتی را گل کند یک متغیر تصادفی برنولی با پارامتر p می‌باشد. بازیکن مربوطه در تمرینات از تعداد ۲۰۰ ضربه پنالتی ۱۶۰ مرتبه را موفق به زدن گل شده است. اگر بخواهیم در مسابقات بازیکن مربوطه را مامور زدن ضربه پنالتی کنیم با چه احتمالی می‌توانیم ۹۵ درصد اطمینان داشته باشیم که وی گل می‌زند؟

حل: معلومات مساله عبارتند از:

$$n = ۲۰۰$$

$$\overline{X} = \hat{p} = \frac{۱۶۰}{۲۰۰} = \frac{۱۶}{۲۰} = ۰/۸$$

$$۱-\alpha = ۰/۹۵ \rightarrow \alpha = ۰/۰۵ \rightarrow Z_{۱-\frac{\alpha}{۲}} = Z_{۰/۹۷۵} = ۱/۹۶$$

بنابراین یک فاصله ۹۵ درصدی برای p عبارتست از:

$$L_۱ = \overline{X} - Z_{۱-\frac{\alpha}{۲}} \left(\sqrt{\frac{\overline{X}(۱-\overline{X})}{n}} \right) = ۰/۸ - ۱/۹۶ \sqrt{\frac{۰/۸\,(۰/۲)}{۲۰۰}} = ۰/۷۴$$

$$L_۲ = \overline{X} + Z_{۱-\frac{\alpha}{۲}} \left(\sqrt{\frac{\overline{X}(۱-\overline{X})}{n}} \right) = ۰/۸ + ۱/۹۶ \sqrt{\frac{۰/۸\,(۰/۲)}{۲۰۰}} = ۰/۸۵$$

بنابراین با ۹۵ درصد اطمینان می‌توانیم بگوییم که احتمال گل شدن ضربه پنالتی توسط بازیکن مربوطه در بازه $(۰/۸۵ , ۰/۷۴)$ قرار دارد.

۵۵ ۱۰ . ۶ . ۵ فاصله اطمینان برای تفاضل نسبت دو جامعه با تعداد نمونه‌های زیاد

اگر دو نمونه با حجم‌های $n_۲ , n_۱$ از دو متغیر تصادفی برنولی در اختیار داشته باشیم، در این صورت برای مقایسه پارامترهای $p_۲ , p_۱$ متغیرهای تصادفی از تفاضل این دو نسبت استفاده می‌کنیم به این ترتیب می‌توان نشان داد که یک فاصله اطمینان $(۱-\alpha)۱۰۰$ درصدی برای تفاضل دو نسبت از دو متغیر تصادفی برنولی عبارتست از:

$$p\left(\overline{X}_۱-\overline{X}_۲-Z_{۱-\frac{\alpha}{۲}} \sqrt{\frac{\overline{X}_۱(۱-\overline{X}_۱)}{n_۱}+\frac{\overline{X}_۲(۱-\overline{X}_۲)}{n_۲}} \le p_۱-p_۲ \le \overline{X}_۱-\overline{X}_۲ + Z_{۱-\frac{\alpha}{۲}} \sqrt{\frac{\overline{X}_۱(۱-\overline{X}_۱)}{n_۱}+\frac{\overline{X}_۲(۱-\overline{X}_۲)}{n_۲}}\right) = ۱-\alpha$$

که در آن $\overline{X}_۲ = \hat{P}_۲ \; , \; \overline{X}_۱ = \hat{P}_۱$

۵۶ مثال ۲۳: در دو شهر A و B احتمال متولد شدن نوزادان پسر را با متغیرهای تصادفی برنولی $X_۲ , X_۱$ نشان می‌دهیم. در شهر A از ۳۰۰ نوزاد متولد شده ۱۸۰ نوزاد پسر می‌باشند و در شهر B از ۴۰۰ نوزاد متولد شده ۱۹۰ نوزاد پسر متولد شده‌اند. می‌خواهیم احتمال متولد شدن نزاد پسر را در دو شهر A و B مقایسه کنیم. برای این منظور از یک فاصله اطمینان ۹۵ درصدی استفاده می‌کنیم مطلوبست حدود فاصله اطمینان؟

حل: معلومات مساله عبارتند از:

$$n_1 = ۳۰۰$$

$$\overline{X}_1 = \hat{p}_1 = \frac{۱۸۰}{۳۰۰} = \frac{۱۸}{۳۰} = ۰/۶$$

$$n_2 = ۴۰۰$$

$$\overline{X}_2 = \hat{p}_2 = \frac{۱۹۰}{۴۰۰} = ۰/۴۷۵$$

$$1-\alpha = ۰/۹۵ \rightarrow \alpha = ۰/۰۵ \rightarrow Z_{1-\frac{\alpha}{2}} = Z_{۰/۹۷۵} = ۱/۹۶$$

$$L_1 = \overline{X}_1 - \overline{X}_2 - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{(1-\overline{X}_1)\,\overline{X}_1}{n_1} + \frac{\overline{X}_2(1-\overline{X}_2)}{n_2}}$$

$$= ۰/۶ - ۰/۴۷۵ - ۱/۹۶ \sqrt{\frac{(۰/۶)۰/۴}{۳۰۰} + \frac{(۰/۴۷۵)۰/۵۲۵}{۴۰۰}} = ۰/۰۵$$

$$L_2 = \overline{X}_1 - \overline{X}_2 + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\overline{X}_1(1-\overline{X}_1)}{n_1} + \frac{\overline{X}_2(1-\overline{X}_2)}{n_2}}$$

$$= ۰/۶ - ۰/۴۷۵ + ۱/۹۶ \sqrt{\frac{(۰/۶)۰/۴}{۳۰۰} + \frac{(۰/۴۷۵)۰/۵۲۵}{۴۰۰}} = ۰/۱۹۸$$

فاصله اطمینان ۹۵٪ $\quad p(۰/۰۵ \le p_1 - p_2 \le ۰/۱۹۸) = ۰/۹۵$

همچنین می‌توان نوشت: $\quad \Rightarrow p(p_2 + ۰/۰۵ \le p_1 - p_2 + ۰/۱۹۸) = ۰/۹۵$

بر اساس مساوی فوق احتمال متولد شدن نوزاد پسر در شهر A با اطمینان ۹۵ درصد در بازه $(p_2 + ۰/۰۵ \ ,\ p_2 + ۰/۱۹۸)$ قرار دارد.

S-۱۴

۵۷ ۱۰ . ۶ . ۶ فاصله اطمینان برای نسبت واریانسهای دو متغیر تصادفی نرمال

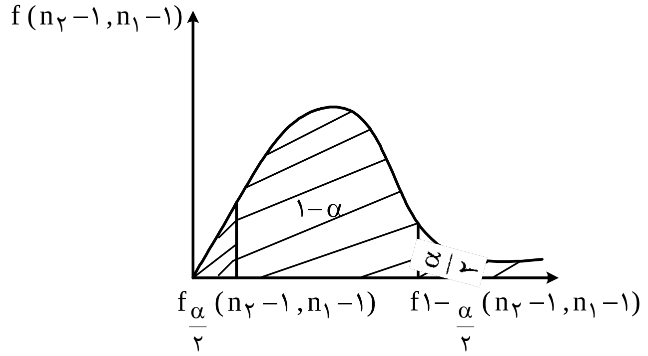
دو نمونه به حجم n_1 و n_2 را با واریانس نمونه‌ای S_1^2 و S_2^2 از دو جامعه نرمال با واریانس σ_1^2 و σ_2^2 استخراج می‌کنیم. برای مقایسه واریانسهای دو جامعه از نسبت آنها $(\frac{S_1^2}{S_2^2})$ استفاده می‌کنیم. قبلاً نشان دادیم که متغیر تصادفی $\frac{S_1^2 \sigma_2^2}{S_2^2 \sigma_1^2}$ دارای توزیع F با $n_2 - 1 \ , \ n_1 - 1$ درجه آزادی است به این ترتیب می‌توان فاصله اطمینان $100(1-\alpha)$ درصدی را برای نسبت $\frac{\sigma_1^2}{\sigma_2^2}$ بصورت زیر بدست آورد:

$$p\left(\frac{S_1^2}{S_2^2}\,\frac{1}{f_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{S_1^2}{S_2^2}\,\frac{1}{f_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)}\right) = 1-\alpha$$

با توجه به اینکه $f_{\frac{\alpha}{2}}(n_1, n_2) = \frac{1}{f_{1-\frac{\alpha}{2}}(n_2, n_1)}$ بنابراین رابطه فوق به صورت زیر ساده می‌شود:

$$p\left(\frac{S_1^2}{S_2^2}\,\frac{1}{f_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{S_1^2}{S_2^2}\, f_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)\right) = 1-\alpha$$

که در آن $f_{\frac{\alpha}{۲}}\ (n_۱-۱, n_۲-۱)$ مقداری از متغیر تصادفی F با $n_۱-۱, n_۲-۱$ درجه آزادی است بطوریکه سطح زیر منحنی به ازای مقادیر کوچکتر از آن برابر $\frac{\alpha}{۲}$ باشد. به شکل زیر توجه کنید:



توجه کنید در تمام فواصل اطمینان که در روابط آنها از $S^۲$ استفاده شده است اگر میانگین جامعه (μ) معلوم باشد به جای استفاده از $S^۲ = \frac{۱}{n-۱}\sum_{i=۱}^{n}(x_i-\overline{X})^۲$ از $S^۲$ استفاده می‌شود که به جای $\overline{X}$ مقدار μ جایگزین شده است یعنی : $S^۲ = \frac{۱}{n-۱}\sum_{i=۱}^{n}(x_i-\mu)^۲$.

۵۸ مثال ۲۴: در یک کارخانه از دو ماشین A و B برای برش میله‌های فولادی استفاده می‌شود در صورتی که واریانس طول میله‌های بریده شده توسط هر یک از ماشینها سه برابر دیگری شود تصمیم به تعمیر ماشین مربوطه می‌گیریم. در یک نمونه ۲۵ تایی از میله‌های بریده شده توسط ماشین A مقدار $S_۱^۲$ برابر ۱۷۰ می‌باشد، همچنین در یک نمونه ۳۱ تایی از میله‌های بریده شده توسط B مقدار $S_۲^۲$ برابر ۲۵ می‌باشد. با اطمینان ۹۵٪ برای نسبت $\frac{\sigma_۱^۲}{\sigma_۲^۲}$ چه تصمیمی می‌توان در مورد ماشین A و B گرفت؟

حل: داریم:

$$n_۱ = ۲۵ \qquad n_۲ = ۳۱$$

$$S_۱^۲ = ۱۷۰ \qquad S_۲^۲ = ۲۵$$

$$۱-\alpha = ۰/۹ \ \rightarrow\ \alpha = ۰/۱ \ \rightarrow\ f_{۱-\frac{\alpha}{۲}}(n_۱-۱\,,\ n_۲-۲) = f_{۰/۹۵}(۲۴\,,\ ۳۰) = ۱/۹۴$$

$$f_{۱-\frac{\alpha}{۲}}(n_۱-۱\,,\ n_۲-۲) = f_{۰/۹۵}(۳۰\,,\ ۲۴) = ۱/۸۹$$

به این ترتیب یک فاصله اطمینان ۹۵ درصدی عبارتست از:

$$p(\frac{۱۷۰}{۲۵}\frac{۱}{۱/۹۴} \le \frac{\sigma_۱^۲}{\sigma_۲^۲} \le \frac{۱۷۰}{۲۵}\,۱/۸۹\ ) = ۰/۹۵$$

$$\Rightarrow p\ (۳/۵ \le \frac{\sigma_۱^۲}{\sigma_۲^۲} \le\ ۱۲/۸۵) = ۰/۹۵$$

بنابراین نسبت واریانسها با ۹۵ درصد اطمینان در بازه $(۱۲/۸۵\ ,\ ۳/۵)$ قرار دارد.

و در نتیجه:

$$p(۳/۵\ ,۱۲/۸۵\ \sigma_۲^۲\ ) = ۰/۹۵$$

یعنی با ۹۵ درصد اطمینان می‌توان گفت که واریانس ماشین A سه برابر واریانس ماشین B می‌باشد بنابراین ماشین A بایستی تعمیر شود.

۵۹ . ۱۰ . ۶ . ۷ فاصله اطمینان برای تفاضل میانگین‌های دو جامعه نرمال

از دو جامعه نرمال با واریانسهای σ_1^2 و σ_2^2 دو نمونه به حجم n_1 و n_2 می‌گیریم. قبلاً نشان دادیم که $Z=\dfrac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$ یک متغیر تصادفی نرمال استاندارد می‌باشد. بنابراین می‌توانیم یک فاصله اطمینان $(1-\alpha)100$ درصد برای تفاضل میانگین‌های دو جامعه نرمال با واریانس معلوم بصورت زیر بدست آوریم:

$$p\left(\overline{X}_1-\overline{X}_2-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\le \mu_1-\mu_2\le \overline{X}_1-\overline{X}_2+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)=1-\alpha$$

و در صورتی که واریانسهای دو جامعه نامعلوم اما مساوی باشند ابتدا واریانس مشترک دو نمونه را بصورت زیر برآورد می‌کنیم:

$$\sigma_1^2=\sigma_1^2=\sigma^2$$

$$\hat{\sigma}^2=S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

فاصله اطمینان در این حالت از رابطه زیر بدست می‌آید:

$$p\left(\overline{X}_1-\overline{X}_2-t_{1-\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\le \mu_1-\mu_2\le \overline{X}_1-\overline{X}_2+t_{1-\frac{\alpha}{2}}(n_1+n_2-2)\,S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)=1-\alpha$$

۶۰ مثال ۲۵: میزان برق مصرفی دو شهر A و B یک متغیر تصادفی نرمال می‌باشد برای مقایسه متوسط مصرف یک نمونه ۱۵ تایی از دو شهر انتخاب می‌کنیم و نتایج زیر را بدست می‌آوریم:

$$\overline{X}_A=2590 \qquad \sigma_A^2=150$$
$$\overline{X}_B=2560 \qquad \sigma_B^2=100$$

مطلوبست:

الف) با فاصله اطمینان ۹۹ درصدی چه نظری می‌توان درباره تفاضل متوسط مصرف در دو شهر داشت؟

ب) اگر برای دو نمونه داشته باشیم $S_A^2=165$, $S_B^2=120$ یک فاصله اطمینان ۹۹ درصدی برای حالتی که واریانسها نامعلوم باشند و مساوی بسازید؟

حل:الف)

$$1-\alpha=0/99 \quad\rightarrow\quad \alpha=0/01 \quad\rightarrow\quad Z_{1-\frac{\alpha}{2}}=Z_{0/99}=2/575$$

$$\rightarrow\quad L_1=2590-2560-2/575\sqrt{\frac{150}{15}+\frac{100}{15}}=19/48$$

$$L_2=2590-2560-2/575\sqrt{\frac{150}{15}+\frac{100}{15}}=40/51$$

بنابراین با ۹۹ درصد اطمینان می‌توانیم بگوییم که $19/48\le\mu_1-\mu_2\le 40/51$ و چون $\mu_1-\mu_2$ در یک بازه مثبت قرار گرفته‌اند می‌توانیم با ۹۹ درصد اطمینان بگوییم که متوسط مصرف شهر A از شهر B بیشتر است.

۶۱ ب)

$$S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}=\frac{14\,(120+165)}{15+15-2}=142/5 \quad\Rightarrow\quad S_p=11/93$$

$$1-\alpha = ۰/۹۹ \quad \rightarrow \quad \alpha = ۰/۰۱ \quad \rightarrow \quad t_{1-\frac{\alpha}{2}}(n_1+n_2-2) = t_{۰/۹۹۵}(۲۸) = ۲/۷۶$$

$$L_1 = \overline{X}_1 - \overline{X}_2 - t_{1-\frac{\alpha}{2}}(n_1+n_2-2)\, S_p \sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = ۲۵۹۰ - ۲۵۶۰ - ۲/۷۶\,(۱۱/۹۳)\sqrt{\frac{1}{۱۵}+\frac{1}{۱۵}} = ۱۷/۹۷$$

$$L_2 = \overline{X}_1 - \overline{X}_2 - t_{1-\frac{\alpha}{2}}(n_1+n_2-2)\, S_p \sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = ۲۵۹۰ - ۲۵۶۰ + ۲/۷۶\,(۱۱/۹۳)\sqrt{\frac{1}{۱۵}+\frac{1}{۱۵}} = ۴۲$$

ملاحظه می‌کنید که در حالت دوم که واریانس مشترک را بصورت تقریبی محاسبه نمودیم باز هم فاصله اطمینان مقداری نزدیک به حالت الف شده است.