تئوری خطاها

جناب آقای دکتر نجفی

تهیه و تنظیم:

گروهی از دوستان مهندسی نقشه برداری دانشگاه زنجان

ورودی سالهای 81 ، 82 ، 83 و 84

گردآوری توسط دوره سوم انجمن علمی گروه نقشه برداری دانشگاه زنجان

مقدمه :

چهار بخش اول تستهای ساده ای از زمینه های مربوطه را که لازم برای درک اساس محاسبات می باشند ارائه می دهند . این بخش ها به تنهائی کامل نیستند و دانشجویان برای تکمیل اطلاعات خود می توانند به منابع دیگر نظیر آنچه که در آخر کتاب لیست شده اند مراجعه کنند .

بخش جداگانه ای (بخش 5) به بیان فلسفی اساس محاسبات سرشکنی اختصاص داده شده است . این بخش ،که نباید آنرا زیاد جامع به حساب آورد ، دلایل استفاده از روش کمترین مربعات در سرشکنی را ارائه می دهد .

سرانجام بخش آخر اصول سرشکنی را بیان می کند .این بخش نیز فقط به مقدمات سرشکنی می پردازد .با علم به اینکه دوره های بعدی برای تکمیل معلومات در این زمینه وجود خواهد داشت .

در نوشته های کتاب تأکید شده است که بحث « آمار گسسته » یعنی آمار نمونه های تصادفی و بحث « آمار پیوسته » یعنی آمار متغیر های تصادفی به موازات هم گسترش یابند . نمونه های تصادفی کمیتهائی هستند که ما در تجربیات روزانه با آنها سروکار داریم در حالیکه ابزار ریاضی یا مدلهای ریاضی مورد استفاده منسوب به فضای پیوسته می باشند .درک صحیح و کامل دو بحث فوق و روابط آنها با همدیگر برای کسی که نخواهد بطور صحیح از محاسبات سرشکنی استفاده کند امری است اجتناب ناپذیر .

فهرست داده شده در آخر کتاب شامل منابع مفیدی در زمینه آمار و سر شکنی می باشد .به خوانندگان علاقمند توصیه می شود برای تکمیل نوشته های این کتاب حداقل به چند منبع ذکر شده در فهرست مراجعه کنند .

1-اساس تئوری مجموعه ها

1-1 تعریف مجموعه(*set*):

مجموعه ای از اشیاء یا عناصر متمایز از همدیگر را یک مجموعه می نامند .

یک مجموعه موقعی مشخص است که تمام عناصر آن مشخص باشد .

مثال1-1

$A_1 \equiv$ {◯ , ☼ , & ,4 , 18}

$A_2 \equiv$ {1, 8,15 , ς , 0 , ☼ ,4 }

$A_3 \equiv$ {0 , 1}

$A_4 \equiv$ {تمام پاهای چپ}

$A_5 \equiv$ {تمام شهرهای با جمعیت بیش از ده میلیون در ایران}

$R \equiv$ {تمام اعداد کسری}

$I \equiv$ {تمام اعداد صحیح مثبت}

در مثال فوق چند مجموعه با اسامی و لیست عناصر تشکیل دهنده آنها آمده است،اگر a یک عنصر یا عنصری از مجموعه A باشد،می گوییم که عنصرa متعلق به مجموعه A است و می نویسیم :

$$a \in A$$

و اگر a متعلق به A نباشد می نویسیم :

$$a \notin A$$

مثال2-1

با مراجعه به مثال 1-1 می بینیم که :

☼$\in A_1$, $8 \in A_2$, $2 \in A_1$, یک پای راست$\notin A_4$

یک قسمت از مجموعه G را شامل یک یا چند عنصر باشد،یک مجموعه فرعی H مـی نامنـد ومی نویسند :

$$H \subset G$$

یعنی H مجموعه ایست متعلق به مجموعه G . اگر مجموعه G نباشد یعنی تمـام عناصـر آن متعلق به G نباشد می نویسیم :

$$H \not\subset G$$

مثال1-3

با مراجعه به مثال 1-1 معلوم می شود که

$A_3 \subset A_2$, $\{2,35\ ,118\} \subset I$, $\{1,8\ ,3\ ,6\ ,2\} \subset R$, $\{\bigcirc\ , \&\ ،\ \varsigma\} \subset A_1$

مجموعه ای را که فاقد عنصری باشد،مجموعه خالی می نامند و با علامت Ø نشان می دهند .

مثال1-4

مجموعه { تمام افراد بلندتر از 3 متر }≡A_6 یک مجموعه خالی است یا بعبارت دیگر $A_6 \equiv$ Ø و همچنین در مثال 1-1 داریم $A_5 \equiv$Ø .

دو مجموعه را مساوی گویند اگر عناصر آنها تک به تک با هم برابر باشند .

مثال1-5

تمام مجموعه های زیر با هم برابرند $\{1,2\ ,3\}$, $\{3,1\ ,2\}$, $\{2,3\ ,1\}$

1-2 تعریف فرایازی (progression) و مجموعه تعاریف (definition set) :

یک ***فرایازی*** ε عبارتست از مجموعه اشیاء مرتب (منظور از مرتب اینست که اشیاء هر کدام مکان مشخصی دارند) که بعضی از آنها قابل تشخیص از همدیگر نیستند .

مجموعه تعاریف D یک فرایازی ε عبارتست از مجموعه متشکل از تمام عناصر قابل تشخیص فرایازی ε . با این تعریف D یک مجموعه فرعی از ε می باشد .

مثال 1-6

برای فرایازی ε

$\varepsilon \equiv (1\ , 2, \bigcirc\ , 2,1\ , 8,\ \&\ )$

مجموعه تعاریف D چنین خواهد بود

$D \equiv \{1\ , 2, \bigcirc\ , 8,\ \&\ \}$

تاکنون باید فرق بین فرایازی و یک مجموعه روشن شده باشد .بطور مثال فرایازی (8, 2,1,1,2, , ♋◯)،

یک فرایازی جدا از فرایازی ذکر شده در مثال 6-1 می باشد . در صورتیکه مجموعه های

{ ◯ , 8, ♋ , 2, 1}   ,   { 2 , 1, 8, ◯ , ♋ } , …

همه با هم مساوی بوده و مجموعه تعاریف فرایازی ε می باشند .

3-1 حاصلضرب کارتزین (Cartesian Product) مجموعه ها :

حاصلضرب کارتزین یا حاصلضرب دو مجموعه A وB عبارتست از یک مجموعه دو تائی که با A×B یعنی A ضربدر B نشان می دهند .هر عنصر از مجموعه دو تائی تشکیل شده از یک زوج عناصر متعلق به A و B که به ترتیب انتخاب شده اند . اگر a یک عنصر متعلق به A ($a \in A$) و b یک عنصر متعلق به B ($b \in B$) باشد ، عنصر دوتائی مجموعه حاصلضرب بصورت (a,b) نشان داده می شود .

$$(a,b) \in A \times B$$

در صورتیکه $b \notin A$ یا $a \notin B$ باشد ، عنصر (b,a) متعلق به A×B نخواهد بود .

$$(b,a) \notin A \times B$$

تعریف حاصلضرب می تواند روی چند مجموعه (بیش از دو) مثلاً n مجموعه تعمیم داده شود . در این حالت مجموعه حاصلضرب یک مجموعه n تائی خواهد بود یعنی هر عنصر دارای n مؤلفه خواهد بود ، بنابراین توان n-ام مجموعه A که با A^n نشان می دهند و برابر است با حاصلضرب مجموعه A به تعداد n بار در خودش .

مثال 7-1

اگر A≡{ 3,1,5} و B≡{ 2,4} حاصلضرب A×B بصورت زیر خواهد بود :

A×B≡{(3,2) , (3,4) , (1,2) , (1,4) , (5,2) , (5,4)}

با مراجعه به مثال 1-1 بسادگی می توان نوشت :

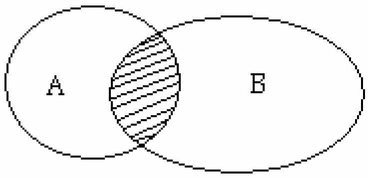
$(4,18, ς) \in A_1 \times A_2$, $(☼,☼) \in A_1 \times A_2$

$(1, ♋) \in A_2 \times A_1$, $(1, ♋) \notin A_1 \times A_2$

$(5,16,3,26,1,0,1) \in R^5$, $(1,2,15,1,8) \in I^5$

4-1 فصل مشترک مجموعه ها (Intersection of sets)

فصل مشترک دو مجموعه A وB که با A∩B نشان داده می شود ، عبارتست از یک مجموعه فرعی از A و B و متشکل از تمام عناصر مشترک یعنی A و B و نه بیـشتر . در شـکل شـماره 1-1 فصل مشترک دو مجموعه A وB با هاشور مشخص شده است .



شکل ۱-۱

اشکالی نظیر شکل فوق را Venn Diagrams می نامند .از شکل فوق بسادگی دیده مـی شـود که

$$A \cap B \equiv B \cap A$$

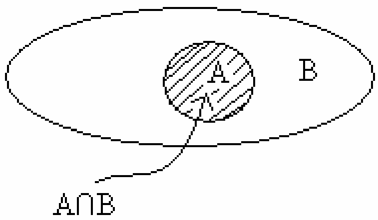
مثال 8-1

با مراجعه به مثال 1-1 دیده می شود

$A_1 \cap A_2 \equiv \{☼\}$, $R \cap I \equiv I$

$A_2 \cap I \equiv \{1,4,8,15\}$, $A_3 \cap A_2 \equiv \{1,0\} \equiv A_3$

اگر مجموعه A یک مجموعه فرعی از مجموعه B باشد ، $A \subset B$، خواهیم داشت $A \cap B \equiv A$ و بر عکس .

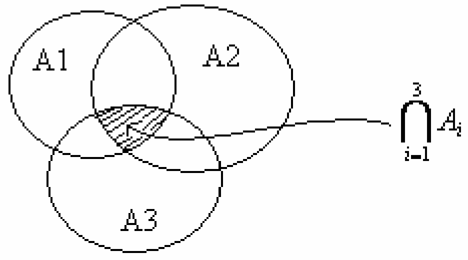


شکل ۲-۱

اگر A∩B≡Ø باشد دو مجموعـه A وB را غیـر متقـاطع (Disjoint Set) مـی نامنـد .فـصل مشترک n مجموعه A_1 وA_2 و. . . و A_n را معمولاً با علامت $\bigcap_{i=1}^{n} A_i$ نشان می دهند .

$$\bigcap_{i=1}^{n} A_i \equiv A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n$$

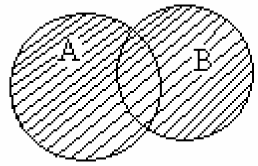
در شکل زیر فصل مشترک سه مجموعه A_1 وA_2 وA_3 نشان داده شده است .



شکل ۳-۱

1-5 انجمن مجموعه ها (Union of set):

انجمن دو مجموعه A وB ،که با $A \cup B$ نشان داده می شود ، عبارت است از مجموعه ای که شامل تمام عناصر A و B و نه بیشتر باشد . انجمن دو مجموعه A وB در شکل با هاشور مشخص شده است .



شکل ۴-۱

انجمن n مجموعه A_1 وA_2 و. . . وA_n را با $\bigcup_{i=1}^{n} A_i$ نشان می دهند

$$\bigcup_{i=1}^{n} A_i \equiv A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n$$

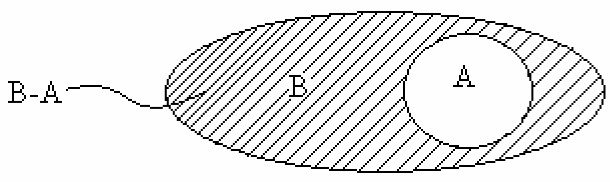
مثال 9-1

با مراجعه به مثال 1-1 می توان نوشت :

$\bigcup_{i=1}^{n} A_i \equiv \{$ ◯ , ☼ , ϫ ,4 , 18,1 , 8,15 , ς , 0,4 $\}$, $I \cup R \equiv R$

با تصور اینکه انجمن دو مجموعه عبارتست از مجموع آن دو ، تفاضل دو مجموعه عبارت خواهد بود از مکمل یکی در داخل دیگری . با در نظر گرفتن دو مجموعه A و B بصورت $A \subset B$ (شکل 5-1)

مجموعه تمام عناصر B صرفنظر از عناصر مشترک بین A و B را مجموعه مکمل A در B می نامند (B-A) .



شکل ۵-۱

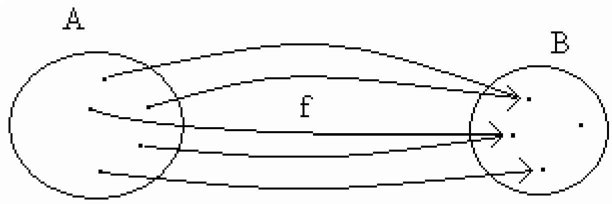
مثال 1-10

در مثال 1-1 مکمل مجموعه A_3 در A_2 عبارتست از :

A_2-A_3≡ {8,15 ,ς , ☼,4} , R-I≡{تمام اعداد کسری بدون اعداد صحیح}

1-6 گسترش مجموعه ها (Mapping of set)

عامل f را فقط یک عنصر از مجموعه B را به یک یا چند عنصر از مجموعه A داخل B می نامند(شکل 6-1) . یعنی برای هر عنصر $a \in A$ فقط یک تصویر $b \in B$ وجود دارد .البته در اینجا منظور ارتباط یک به یک یا تناظر دو سوئی (یعنی برای هر یک عنصر $b \in B$ نیز یک عنصر $a \in A$ وجود دارد) مورد نظر نیست .



شکل ۶-۱

گسترش f را بصورت زیر نشان می دهیم :

$$f \in \{A \rightarrow B\}$$

یعنی f عنصری یا عضوی از مجموعه تمام گسترش های A داخل B یا بعبارت ساده f یک گسترش A داخل B و یا f ، A را داخل B تصویر می کند . اگر عناصر مجموعه B کلاً تصاویر عناصر مجموعه A باشند ، f را یک گسترش A روی B می نامند و یا f ، A را روی B تصویر می کند .

اگر A و B دو مجموعه عددی باشند f را یک تابع (یک رابطه ریاضی بین $a \in A$ و تصویر مربوطه اش $b \in B$) می نامند . در این صورت b برابر خواهد بود با مقدار تابع بازای a یعنی :

$$b = f(a)$$

مثال 1-11

مجموعه $A \equiv \{a_1, a_2, a_3\} \equiv \{2,-1,3\}$ و گسترش $f \in \{A \to B\}$ بطوریکه $f(a_i) = a_i^3$ را برای هر $a_i \in A$ داده شده است . عناصر مجموعه B بصورت زیر محاسبه می شود :

$$b_i = f(a_i) = a_i^3$$

$b_1=(2)^3= 8$, $b_2=(-1)^3= -1$, $b_3=(3)^3 =27$

عموماً f یک گسترش A داخل B است یعنی $\{8,-1,27\} \subset B$ است اگر f یک گسترش A روی B باشد خواهیم داشت : $\{8,-1,27\} \equiv B$

1-7 تمرینات

1- کدام یک از مجموعه های زیر با هم برابرند ؟

$\{t , r , s\}$, $\{s , r , t\}$, $\{r , s , t\}$, $\{t , s , r\}$

2- با در نظر گرفتن مجموعه های

$A\equiv\{d\}$, $B\equiv\{c , d\}$, $C\equiv\{a , b , c\}$, $D\equiv\{a , b\}$, $H\equiv\{a , b , d\}$

صحت یا سقم روابط زیر را معلوم کنید .

1) $B \subset D$ 2) $D \subset C$ 3) $A \subset H$ 4) $(A \cap B) \not\subset C$

5) $C\equiv B$ 6) $B \neq H$ 7) $(A \cup D) \subset H$ 8) $(H \cap C) \equiv D$

3- با در نظر مجموعه های زیر

$U\equiv\{ 1,2 ,3 ,. . . ,8 ,9 \}$, $A\equiv\{1,2 ,3 ,4 \}$, $B\equiv\{2,4 ,6 ,8 \}$

$C\equiv\{3,4 ,5 ,6 \}$, $D\equiv\{1,3,5,7,9 \}$

حاصل روابط زیر را بدست آورید .

1) $B \cup D$ 2) $A \cap C$ 3) $A \cup B$ 4) $U - A$

5) مجموعه H را چنان پیدا کنید که یک مجموعه فرعی از مجموعه های U و A و D باشد .

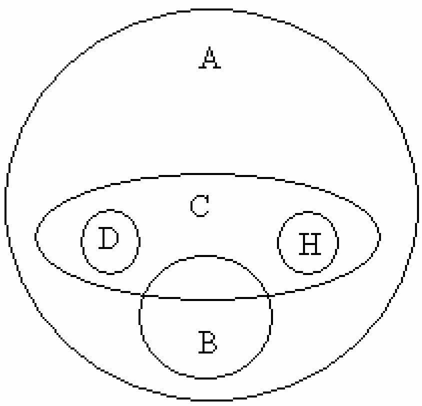
4- با در نظر گرفتن مجموعه های A و B و C و D و H در شکل زیر حاصل روابط زیر را با هاشور روی شکل مشخص نمایید .

1) $D \cup H$

2) $H \cap C$

3) $C \cap B$

4) $A - C$



شکل ۷-۱

5) $B \cup C$

6) $(A-B) \cup (C \cap B)$

7) $A-(C \cup B)$

5- با در نظر گرفتن دو مجموعه زیر

$A \equiv \{3,4,0,-1\}$, $B \equiv \{-2,5\}$

حاصلضرب $A \times B$ و $B \times A$ را پیدا کنید و همچنین توان دوم B یعنی B^2 را محاسبه کنید .

6- مجموعه $X \equiv \{-2,-1,0,1,2\}$ و گسترش $f \in \{x \to y\}$ بطوریکه برای هر $x \in X$ تابع $f(x) = x2+1$ داده شده است . مجموعه Y را با توجه با اینکه f یک گسترش X روی Y می باشد پیدا کنید.

2-اساس تئوری ریاضی احتمالات

2-1 فضای احتمال ، تابع احتمال و احتمال

مجموعه $D \neq \phi$ را در نظر می گیریم .فرض کنیم که این مجموعه میتواند به مجموعه های فرعی غیر متقاطع D_j ، $D_j \subset D$ ، بطوریکه $D \equiv \bigcup_j D_j$ تقسیم شود (منظور از مجموعه های غیر متقاطع یعنی $D_i \bigcap D_j \equiv \phi$ برای هر $i \neq j$) مجموعه D را یک فضای احتمال می نامیم .

هر نوع گسترش P مجموعه D را روی مجموعه $[0,1]$ (مجموعه تمام اعداد حقیقی مثبت بین دو عدد صفر و یک $0 \leq b \leq 1$) که دارای دو خاصیت زیر باشد تابع احتمال نامیده می شود :

- اگر $D' \subset D$ باشد، $P(D') = 1 - P(D - D')$ می باشد ($D - D'$ عبارتست از مکمل D' در D)
- اگر مجموعه های فرعی D_i متعلق به D بوده ، یعنی $D_1, D_2, \ldots, D_n \subset D$ ، و غیر متقاطع باشند داریم $P(\bigcup_{i=1}^{n} D_i) = \sum_{i=1}^{n} P(D_i)$.

مقدار تابع احتمال ، $P(D')$ ، را که عددی در فاصله $[0,1]$ می باشد احتمال می نامند . از دو خاصیت ذکر شده برای تابع احتمال در بالا ،نتایج زیر گرفته می شود :

1) $P(D)=1$
2) $P(\emptyset)=0$
3) اگر $D' \subset D$ باشد داریم $P(D') \leq 1$
4) اگر $D'' \subset D'$ باشد داریم $P(D'') \leq P(D')$
5) اگر $A, B \subset D$ و $A \bigcap B \equiv \phi$ باشد داریم $P(A \bigcup B) = P(A) + P(B)$

اگر D مجموعه ای از نقاط (Point Set) باشد ، یعنی عناصر آن بتوانند بوسیله نقاط معرفی شوند ، آن مجموعه همیشه قابل تفکیک می باشد . مقدار $\sum P(D_i) \in [0,1]$ را گاهی احتمال کل یا احتمال مجموع (accumulative Probability) $\bigcup_i D_i$ می نامند .

2-2 احتمال شرطی (Conditional Probability)

اگر دو مجموعه A و B متعلق به مجموعه D باشند ، $A, B \subset D$ ، نسبت

$$P(A \bigcap B)/P(B) = P(A|B)$$

را احتمال شرطی می نامند .طرف راست رابطه فوق یعنی $P(A|B)$ را احتمال A وقتی B داده شده است می خوانند .بعبارت دیگر احتمال شرطی $P(A|B)$ یعنی احتمال وقوع A مشروط بر آنکه B اتفاق افتاده باشد.

از تعریف احتمال شرطی مشاهده می شود که :

1.اگر $P(B)=0$ باشد ، $P(A|B)$ نامعین است .

2.اگر $B \subset A$ باشد داریم $A \cap B \equiv B$ و در نتیجه $P(A|B)=1$.

3.اگر $A \cap B \equiv \emptyset$ باشد یعنی A وB غیر متقاطع باشند خواهیم داشت $P(A|B)=0$.

3-2 احتمال مرکب (Combined Probability)

اگر احتمال شرطی $P(A|B)$ مساوی $P(A)$ باشد معلوم است که وقوع A بستگی به وقوع B ندارد.

در چنین حالتی می گوییم A وB مستقل از همدیگر هستند. با استفاده از تعریف احتمال شرطی می توان نوشت:

$$P(A \cap B) = P(A).P(B)$$

رابطه فوق را می توان چنین تعبیر کرد که احتمال وقوع همزمان A وB که معمولاً با $P(A,B)$ نشان میدهند،

خوانده می شود احتمال A وB ، و به احتمال مرکب A وB معروف است ، برابر است احتمال A ضربدر احتمال B .

$$P(A,B) = P(A).P(B)$$

همینطور احتمال وقوع همزمان مجموعه های $D_1, D_2, \ldots, D_n \subset D$ برابر است با حاصلضرب احتمال هر کدام از آنها یعنی :

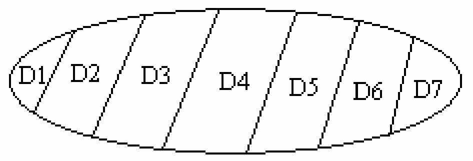
$$P(D_i, D_j) = P(D_i) . P(D_j) \qquad i \neq j$$

$$P(D_i, D D_k) = P(D_i) . P(D_j) . P(D_k) \qquad i \neq j \quad , \quad i \neq k$$

$$P(D_1, D_2, \ldots, D_n) = \prod_{i=1}^{n} P(D_i)$$

مثال 1-2

فرض کنید که فضای احتمال D به هفت مجموعه فرعی غیر متقاطع $D_1, \ldots, D_7$ تقسیم شده است بطوریکه $D = \bigcup_{i=1}^{7} D_i$



شکل ۱-۲

فرض می کنیم که احتمال مجموعه ها، $P(D_i)$ ها ، محاسبه شده و معلوم می باشند .

$$P(D_1)=\frac{1}{28} \quad , \quad P(D_2)=\frac{2}{28} \quad , \quad P(D_3)=\frac{3}{28} \quad , \quad P(D_4)=\frac{4}{28}$$
$$P(D_5)=\frac{5}{28} \quad , \quad P(D_6)=\frac{6}{28} \quad , \quad P(D_7)=\frac{7}{28}$$

احتمال کل مجموعه ها برابر است با :

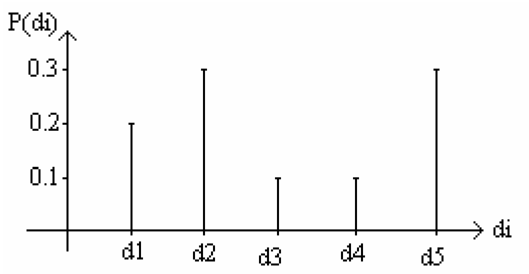
$$P(D)=P(\bigcup_{i=1}^{7} D_i)= \sum_{i=1}^{7} P(D_i)=\frac{1+2+3+4+5+6+7}{28}=1$$

احتمال مرکب تمام D_i ها برابر است با :

$$P(D_1,D_2,\ldots,D_7)=\prod_{i=1}^{7} P(D_i)=0$$

مثال 2-2

در این مثال فرض می کنیم که فضای احتمال D به 5 نقطه d_i تقسیم شده است $j=1,2,\ldots,5$ ، $d_i \in D$ ، فرض می کنیم احتمال های $P(d_i)$ برای هر نقطه مشخص بوده و روی محور قائم در شکل زیر آورده شده اند .



شکل ۲-۲

$P(d_1)=0.2 \quad , \quad P(d_2)=0.3 \quad , \quad P(d_3)=0.1 \quad , \quad P(d_4)=0.1 \quad ,$
$P(d_5)=0.3$

احتمال کل $P(D)=P(\bigcup_j d_j)=\sum_{j=1}^{5} P(d_j)=0.2+0.3+0.1+0.1+0.3=1$

بطور مثال احتمال مرکب d_1 وd_2 برابر است با

$$P(d_1,d_2)=\prod_{j=1}^{2} P(d_j) = P(d_1)\;.\;P(d_2)=0.2\times 0.3=0.06$$

این احتمال مرکـب عـبارتست از احتمال وقوع همزمان d_1 و d_2 با فرض اینکه d_1 و d_2 مستقـل از همدیگر می باشند .

2-4 تمرین شماره 2

هر شماره روی تاس یک احتمال ظهور دارد وقتی که تاس ریخته میشود مجموعه های {اعداد زوج }≡A و{ اعداد اول}≡B و{ اعداد فرد}≡C عبارتند از مجموعه های فرعی از مجموعه تمام اعداد روی تاس .

1. فضای احتمال D را برای شماره های تاس تعیین کنید .
2. احتمال وقوع یا ظهور هر شماره از تاس را معین کنید .
3. احتمال های P(A) وP(B) وP(C) را پیدا کنید .
4. احتمال اینکه یک عدد زوج و یا عدد اول ظاهر شود چقدر است .
5. احتمال اینکه یک عدد فرد ظاهر شود چقدر است .

3-اساس آمار (Fundamentals of Statistics)

3-1 آمار یک نمونه واقعی (Actual Sample)

3-1-1 تعریف یک نمونه تصادفی (Random Sample)

هر فرایازی محدود (شامل تعداد معین و محدود n از عناصر) ε

$$\varepsilon \equiv (\varepsilon_1 , \varepsilon_2 , \ldots , \varepsilon_n)$$

بطوریکه :

أ- مجموعه تعاریف آن ، یعنی D ، را بتوان یک فضای احتمال تعریف کرد

ب-دارای تابع احتمال P که برای هر عنصر $d_i \in D$ بصورت $P(d_i) = \frac{C_i}{n}$ تعریف شده باشد (C_i عبارتست از شماره یا فرکانس عنصر d_i در فرایازی ε) یک نمونه تصادفی نامیده می شود . نسبت $\frac{C_i}{n}$ را فرکانس نسبی می نامند .

مثال 3-1 فرایازی زیر را در نظر می گیریم

$$\begin{array}{ccccccccc} & \varepsilon_1 & \varepsilon_2 & \varepsilon_3 & \varepsilon_4 & \varepsilon_5 & \varepsilon_6 & \varepsilon_7 & \\ \varepsilon \equiv (& 1, & \&, & \bigcirc, & 1, & 1, & \varsigma, & \& &) \end{array}$$

ε دارای 7 عنصر می باشد (n=7) . مجموعه تعاریف D بصورت زیر خواهد بود :

$$\begin{array}{cccccc} & d_1 & d_2 & d_3 & d_4 & \\ D \equiv \{ & 1, & \&, & \bigcirc, & \varsigma & \} \end{array}$$

مجموعه D از چهار عنصرتشکیل شده است (m=4) و فرکانس های مربوطه عبارتند از

$$C_1=3 \quad , \quad C_2=2 \quad , \quad C_3=1 \quad , \quad C_4=1$$

احتمال های مربوطه یا فرکانس های نسبی عبارت خواهد بود از :

$$P(d_1)=P(1)= \frac{C_1}{n} = \frac{3}{7} \qquad , \qquad P(d_2)=P(\&)=\frac{C_2}{n} = \frac{2}{7}$$

$$P(d_3)=P(\bigcirc)=\frac{1}{7} \qquad , \qquad P(d_4)=P(\varsigma)=\frac{1}{7}$$

ملاحظه می شود که تابع احتمال تعریف شده درفوق دارای خواص یک تابع احتمال می باشد .

بخصوص احتمال کل ، با استفاده از احتمالات فوق چنین محاسبه می شود

$$P(D) = P(\bigcup_{i=1}^{4} d_i) = \sum_{i=1}^{4} P(d_i) = \frac{2}{7} + \frac{2}{7} + \frac{1}{7} + \frac{2}{7} = 1$$

بنابراین هر فرایازی محدود از عناصر ممکن است یک نمونه تصادفی تلقی شود .

این یک کشف مهمی است که باید در موارد زیر مورد نظر قرار گیـرد . در نتیجـه همیـشه مـی توان یک فضای احتمال با احتمالهای مربوطه (احتمال های مربوط به عناصر مجموعـه تعـاریف) را از یک نمونه ساخت .

از حالا به بعدما با فرایازی های عددی ، در نتیجه مجموعه تعاریف $D \subset R$ عـددی ، سـروکار خواهیم داشت. همچنین مجموعه D که در آن اعـداد بـه ترتیـب صـعودی یـا نزولـی معمـولاً صعودی مورد نظر است . ضمناً باید گفت که تعریف ماازیک نمونـه تـصادفی ، در اینجـا ، یـک تعریف استاندارد بدین معنی که یک نمونه تصادفی شامل عناصر بسیار زیادی است،نیـست . در این مورد در بخش آینده شد.

مثال 2-3 یک تاس صد بار ریخته می شود . جدول زیر هر شماره روی تاس را با فرکـانس یـا تعداد دفعاتی که ظاهر می شود نشان می دهد .

شماره	d_i	1	2	3	4	5	6
تعداد	C_i	14	17	20	18	15	16

احتمالهای زیر را پیدا کنید

1. احتمال اینکه عدد 3 ظاهر می شود.
2. احتمال اینکه عدد 5 ظاهر می شود.
3. احتمال اینکه یک شماره زوج ظاهر می شود.
4. احتمال اینکه یک عدد اول ظاهر می شود.

راه حل :

$$1-\ P(3)=\frac{20}{100}=0.2$$

$$2-\ P(5)=\frac{15}{100}=0.15$$

$$3-\ P(2,4,6)=P(2)+P(4)+P(6)=\frac{17}{100}+\frac{18}{100}+\frac{16}{100}=0.51$$

$$4-\ P(2,3,5)=P(2)+P(3)+P(5)=\frac{17}{100}+\frac{20}{100}+\frac{15}{100}=0.52$$

3-1-2 تابع پخش احتمال PDF واقعی (تجربی) و تابع پخش تجمعی CDF

یک نمونه تصادفی ε را که عبارتند از یک فرایازی عددی و در نتیجه مجموعه تعاریف D یک مجموعه عددی است در نظر می گیریم . تابع پخش احتمال P عبارتست از یک تابع گسسته که D را روی [0,1] تصویر می کند . این تابع را معمولاً تابع پخش تجربی (واقعی) یا تابع فرکانس

تجربی نمونه ε می نامند و با علامت اختصاری PDF نشان می دهند . مقدار $P(d_i)$ ، $d_i \in D$ را احتمال تجربی یا فرکانس نسبی عنصر d_i می نامند .

مثال 3-3

نمونه تصافی زیر را که حاصل یک عمل تجربی است در نظر می گیریم

$\varepsilon \equiv (1 , 2 , 4 , 1 , 1 , 2 , 1 , 1 , 2) \quad , \quad n=9$

مجموعه D به صورت زیر خواهد بود

$D \equiv \{1 , 2 , 4\} = \{d_i \;,\; i=1,2,3\} \quad , \quad m=3$

فرکانس های C_i برای هر d_i عبارتند از

$C_1=5 \quad , \quad C_2=3 \quad , \quad C_4=1$

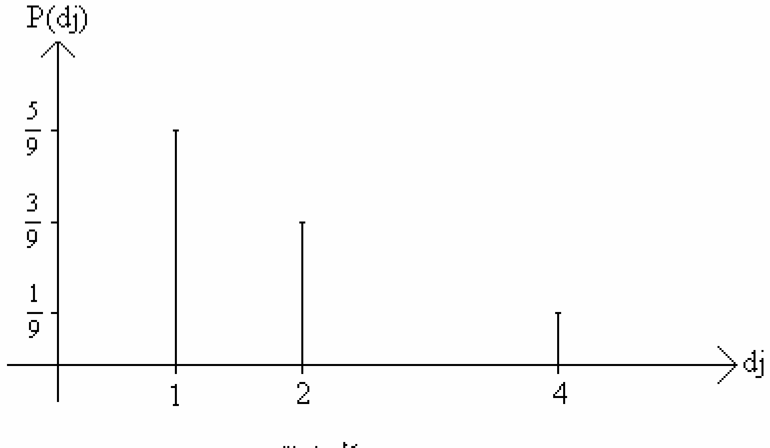
احتمالهای مربوطه برابرند با

$P(1)=\frac{5}{9} \quad , \quad P(2)=\frac{3}{9} \quad , \quad P(4)=\frac{1}{9}$

مقادیر فوق را می توان با رابطه زیر امتحان کرد

$$\sum_{i=1}^{3} P(d_i) = \frac{5}{9}+\frac{3}{9}+\frac{1}{9}=1$$

دیاگرام نمایش تابع احتمال PDF نمونه ε در شکل زیر نشان داده شده است بطوریکه محور افقی نشان دهنده ، d_i ها و محور عمودی نشان دهنده ، $P(d_i)$ ها می باشد . این نمایش را Bar diagram نیز می نامند .



شکل ۱-۳

از آنجاییکه ما از مجموعه های عددی مرتب شده (صعودی) D استفاده می کنیم --- مورد نیست اگر سؤال شود مثلا احتمال اینکه d در فاصله $D' \subset D$ ، $D' \equiv \{d_k , d_j\}$ باشد چقدر

است . چنین احتمالی با $P(D')$ و یا $P(d_k \le d \le d_j)$ نشان داده می شود . بـرای جـواب دادن به این سؤال از تابع پخش احتمال تجربی استفاده می کنیم .

$$P(d_k \le d \le d_j) = \sum_{i=k}^{j} P(d_i) \qquad (3\text{-}1)$$

رابطه فوق احتمالی را که d متعلق به مجموعه D' باشـد ($d \in D' \equiv \{d_k, \ldots, d_j\} \subset D$) نـشان می دهد نه احتمالی را که d متعلق به فاصله $[d_k , d_j]$ باشد .بعبارت دیگر رابطه فوق احتمالی را که d یکی از مقادیر $d_k, d_{k+1}, \ldots, d_{j-1}, d_j$ را انتخاب کند نشان می دهد نه احتمال ایـن را کـه d هر مقدار دلخواهی در فاصله پیوسته $[d_k , d_j]$ اختیار کند .

تابع C از $d_i \in D$ را که شکل زیر داده می شود تابع پخش تجمعی (CDF) تجربی نمونه ε مـی نامند .

$$C(d_i) = \sum_{j \le i} P(d_i) \in [0,1] \qquad (3\text{-}2)$$

مثال 4-3

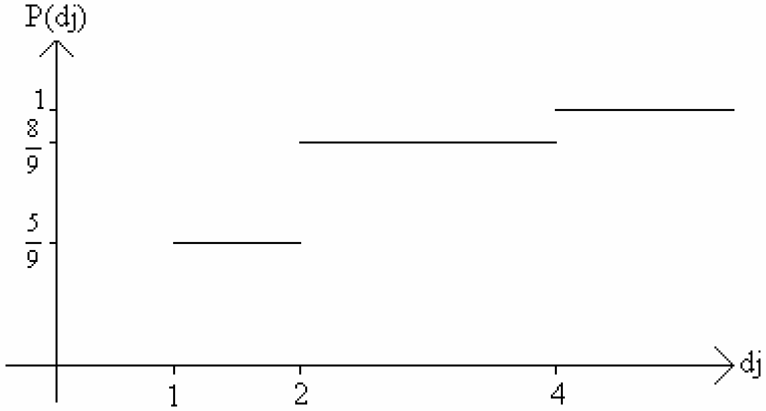
در مثال 3-3 می توان تابع CDF را بصورت زیر محاسبه کرد

$C(d_1)=P(d_1)=\frac{5}{9}$

$C(d_2)=P(d_1)+P(d_2)=\frac{5}{9}+\frac{3}{9}=\frac{8}{9}$

$C(d_3)= P(d_1)+P(d_2)+P(d_3)=\frac{5}{9}+\frac{3}{9}+\frac{1}{9}=1$

در شکل زیر منحنی نمایش تغییرات گسسته CDF نسبت به d_i نشان داده شده است .



شکل ۲-۳

مشاهده می شود که تابع CDF دارای خواص زیر است :

1) مقدار تابع CDF همیشه مثبت است .

2) تابع CDF یک تابع صعودی است .

3) مقدار تابع CDF به ازای بزرگترین d_i یعنی d_m برابر واحد است $C(d_m)=1$
آخرین عضو تابع CDF احتمال کل را نشان می دهد .

مثال 5-3

در مثال 2-3 می توان تابع CDF را بصورت زیر محاسبه کرد

$C(1)=P(1)=0.14$ شماره یک روی تاس

$C(2)=CP(1)+P(2)=0.14 +0.17=0.31$ شماره 2روی تاس

$C(3)= C(2)+P(3)= 0.31+0.20=0.51$ شماره 3 روی تاس

$C(4)=C(3)+P(4)=0.51 +0.18=0.69$ شماره 4 روی تاس

$C(5)=C(4)+P(5)=0.69 +0.15=0.84$ شماره 5 روی تاس

$C(6)=C(5)+P(6)=0.84 +0.16=1$ شماره 6 روی تاس

ملاحظه می شود که بیشترین مقدار CDF یعنی $C(6)$ مساوی واحد می باشد .

منحنی نمایش تغییرات CDF را می توان نظیر مثال قبل ترسیم نمود .

3-1-3 میانگین یک نمونه (Mean of a sample)

فرایازی ε و مجموعه تعاریف D آنرا در نظر می گیریم

$$\varepsilon \equiv (\varepsilon_1 , \varepsilon_2 , \ldots , \varepsilon_n) \quad , \quad D\equiv\{d_1,d_2,\ldots,d_m\}$$

عدد کسری M تعریف شده بصورت زیر را میانگین یا متوسط نمونه ε می نامند .

$$M = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i \in [d_1, d_m] \quad , \quad (3\text{-}3)$$

می توان نشان داد که M از رابطه زیر نیز محاسبه می شود

$$M = \sum_{i=1}^{m} d_i P(d_i) \quad , \quad (3\text{-}4)$$

رابطه فوق را می توان چنین اثبات کرد :

$$\sum_{i=1}^{m} d_i P(d_i) = \sum_{i=1}^{m} d_i (\frac{C_i}{n}) = \frac{1}{n}\sum_{i=1}^{m} C_i d_i = \frac{1}{\mathrm{n}}(\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n) = \frac{1}{n}\sum_{i=1}^{m}\varepsilon_i = M$$

میانگین یک نمونه را می توان چنین تعبیر نمود که عبارتست از نتیجه اعمال اپراتور جمع($\sum$) روی نمونه تقسیم بر تعداد عناصر نمونه (m) و اغلب باعلامت زیر نشان داده می شود.

$M=E(\varepsilon)=\varepsilon$ میانگین $=\varepsilon$ متوسط $=\bar{\varepsilon}$, (3-5)

علامت E مخفف (Mathematical Expectation) و امید ریاضی خوانده می شود . باید در نظر گرفت که امید ریاضی یک اسم دیگر برای علامت $\sum$ می باشد وقتی که روی $d_iP(d_i)$ اعمال می شود (فرمول 4-3) .

E یک اپراتور خطی است و دارای خواص زیر می باشد .

1-E(k)=k
2-E(kε)=kE(ε)
3-E(k+ε)=k+E(ε)
4-$E((\sum_j \varepsilon^j) = \sum_j E(\varepsilon^j)$

ε^j بازای $j=1,2,\ldots,s$ عبارتند از تعداد S نمونه جداگانه با تعداد مساوی m عناصر در مجموعه تعاریف شان D^j. ε^j نباید با ε_j اشتباه شود اولی نشان دهنده یک نمونه است در صورتیکه دومی نشان دهنده عنصری در یک نمونه است بعبارت دیگر ε_j یک عنصر واحد در یک نمونه است در صورتیکه ε^j یک عدد است داخل یک سری نمونه ها .اگر(ε_1) =ε

5-$E(\varepsilon)=\varepsilon_1$
5-$E(E(\varepsilon))=E(\varepsilon)$

مثال 6-3

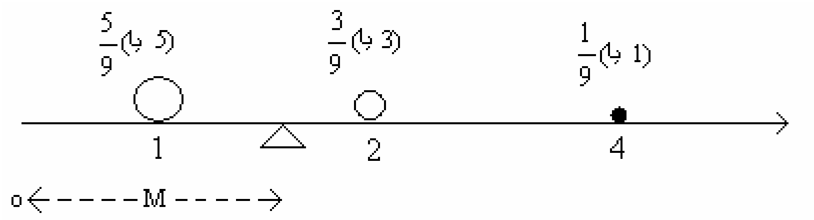
میانگین نمونه داده شده در مثال 3-3 را با استفاده از فرمول (3-3) و(4-3) بدست آورید .

$$M=E(\varepsilon)=\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i=\frac{1}{9}(1+2+4+1+1+2+1+1+2)=\frac{15}{9}=1\frac{2}{3}$$

$$M=E(\varepsilon)=\sum_{j=1}^{m}d_iP(d_i)=1\times\frac{5}{9}+2\times\frac{3}{9}+4\times\frac{1}{9}=\frac{1}{9}(5+6+4)=\frac{15}{9}=1\frac{2}{3}$$ امید ریاضی

واضح است که دو فرمول (3-3) و (4-3) جواب یکسانی خواهند داد .

جالب است که محاسبه میانگین یک نمونه تصادفی با استفاده از (4-3) مشابه است با محاسبه مرکز تعادل در مکانیک . این تشابه به سهولت دیده می شود اگر احتمال های $P(d_i)$ شمارش های C_i را بمثابه وزن ها در مکانیک تلقی کنیم . در شکل 3-3 که مقادیر موجود در مثال 3-3 را مورد استفاده قرار داده است مرکز تعادل نسبت به یک نقطه مبدأ o با مساوی صفر قرار داده مجموع ممانها نسبت به نقطه o محاسبه می شود .



شکل ۳-۳

فاصله بدست آمده برای نقطه تعادل نسبت به نقطه o چیزی نیست جز میانگین نمونه . بر اساس مشابهت فوق با مکانیک ، میانگین محاسبه شده از فرمول 4-3 را میانگین وزن دار (Weighted Mean) نیز می گویند که در آن هر عنصر $d_i \in D$ باندازه احتمال خودش یعنی $P(d_i)$ وزین شده است .

3-1-4 واریانس یک نمونه (Variance of a Sample)

نمونه واقعی ε و مقدار میانگین آن M را در نظر می گیریم

$$\varepsilon \equiv (\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n)$$

عدد کسری S^2 تعریف شده شکل زیر را واریانس (Dispersion) نمونه می نامند .

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(\varepsilon_i - M)^2$$

جذر واریانس یعنی S را **Standard deviation** (انحراف معیار) نمونه می نامند.

با در نظر گرفتن رابطه بین نمونه ε و مجموعه D می توان نوشت

$$\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 = \sum_{j=1}^{m} d_j^2 P(d_j)$$

با استفاده از رابطه فوق می توان فرمول دیگری برای S^2 بدست آورد :

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(\varepsilon_i - M)^2 = \sum_{j=1}^{m} P(d_j)(d_j - M)^2 \qquad (3\text{-}7)$$

می توان گفت که واریانس عبارتست از حاصل اعمال اپراتور E روی $(\varepsilon - E(\varepsilon))^2$ و یا $P(d_j)(d_j - M)^2$

می باشد و معمولاً بصورت زیر نوشته می شود

$$S^2 = E\left((\varepsilon - E(\varepsilon))^2\right) = Var(\varepsilon) \qquad (3\text{-}8\text{ a})$$

رابطه فوق به صورت زیر بسط داده می شود :

$$S^2 = E\left(\varepsilon^2 - 2\varepsilon E(\varepsilon) + E^2(\varepsilon)\right) = E(\varepsilon^2) - 2E(\varepsilon)E(\varepsilon) + E^2(\varepsilon)$$
$$= E(\varepsilon^2) - E^2(\varepsilon) = E(\varepsilon^2) - M^2 \qquad (3\text{-}8\ b)$$

در محاسبات فوق از خواص خطی بودن اپراتور E استفاده شده است و همچنین $E(\varepsilon)=M$ قـرار داده شده است. در نتیجه واریانس بصورت زیر در می آید :

$$S^2 = \sum_{j=1}^{m} d_j P(d_j) - M^2 \qquad (3\text{-}9)$$

با توجه بوجود تشابه با مکانیک می توان واریانس یک نمونه را ،(3-7) مثل ممان اینرسی سیستم نقاط وزن دار (شکل 3-3) نسبت به M دانست .

مثال 7-3

واریانس S^2 نمونه ε داده شده در مثال 3-3 را با استفاده از معادله 3-8 b محاسبه می کنیم

$$E(\varepsilon^2) = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 = \frac{1}{9}(1+4+16+1+1+4+1+1+4) = \frac{33}{9}$$

مقدار فوق را در معادله 3-8b قرار داده و با دانستن $M = \frac{15}{9}$ از مثال 3-6 واریـانس بـصورت زیر محاسبه

$$S^2 = Var(\varepsilon) = \frac{33}{9} - (\frac{15}{9})^2 = \frac{8}{9} = 0.89$$

می شود .

جذر مقدار فوق را انحراف معیار (Standard deviation) نمونه ε می نامیم .

$$S = \sqrt{\frac{8}{9}} = \frac{2\sqrt{2}}{3} = 0.943$$

برای محاسبه واریانس از معادله 3-9 نیز می توان استفاده کرد که همـان نتـایج فـوق را خواهـد داد.

$$\sum_{j=1}^{m} d_j^2 P(d_j) = 1\times\frac{5}{9} + 4\times\frac{3}{9} + 1\times\frac{1}{9} = \frac{1}{9}(5+12+16) = \frac{33}{9}$$

با قرار دادن $M = \frac{15}{9}$ خواهیم داشت :

$$S^2 = \frac{33}{9} - (\frac{15}{9})^2 = \frac{8}{9} = 0.89$$

باید گفته شود که معادلات 3-6 و 3-7 نیز همان نتایج فوق را برای واریانس نمونه خواهد داد . در هر صورت از نقطه نظر محاسبات ،مخصوصاً برای نمونه های ------- ، معادله 3-9 نسبت

به سایرین اولویت دارد . همینطور در محاسبه میانگین یک نمونـه ، بهتـر اسـت از معادلـه 3-4 استفاده شود .

3-1-5 سایر مشخصات یک نمونه : میانه و عرض (Median and Range)

میانه یک نمونه ε ،$\varepsilon \equiv (\varepsilon_1 , \varepsilon_2 , \ldots , \varepsilon_n)$ ، با Med(ε) نشان داده می شود ، به دو صورت برای n فرد وn زوج تعریف می شود . برای n فرد ، میانه عبارتست از عنصر وسطی در فرایازی مرتب شده ε

$$Med(\varepsilon) = \varepsilon_{(\frac{n+1}{2})} \qquad (3\text{-}10)$$

برای n زوج ، میانه عبارتست از میانگین دو عنصر $\varepsilon_{\frac{n}{2}+1}, \varepsilon_{\frac{n}{2}}$

$$Med(\varepsilon) = \frac{1}{2}\left(\varepsilon_{\frac{n}{2}+1} + \varepsilon_{\frac{n}{2}}\right) \qquad (3\text{-}11)$$

مثال 3-8

نمونه $\varepsilon \equiv (5, 3, 6, 4, 1, 2)$ را در نظر می گیریم .برای تعیین Med(ε) ، اول نمونه را خواه در جهت صعودی و خواه نزولی مرتب می کنیم مـثلاً $\varepsilon \equiv (1, 2, 3, 4, 5, 6)$ کـه در آن n=6 مـی باشد . از آنجائیکه n زوج است خواهیم داشت

$$Med(\varepsilon) = \frac{1}{2}\left(\varepsilon_{\frac{n}{2}+1} + \varepsilon_{\frac{n}{2}}\right) = \frac{1}{2}(\varepsilon_3 + \varepsilon_4) = \frac{1}{2}(3+4) = 3.5$$

همینطور میانه نمونه داده شده در مثـال 3-3 راپـس از مرتـب کـردن آن در جهـت صـعودی بدست می آوریم

$$\varepsilon \equiv (1 , 1 , 1 , 1 , 1 , 2 , 2 , 2 , 4) \quad n=9$$

در اینجا n فرد است پس داریم :

$$Med(\varepsilon) = \varepsilon_{\frac{n+1}{2}} = \varepsilon_5 = 1$$

عرض (Range) نمونه $\varepsilon \equiv (\varepsilon_i , i = 1,\cdots,n)$ عبارتست از اختلاف بین بزرگترین عنـصر (ε_l) و کوچکترین عنصر (ε_S) در نمونه .

$$Ra(\varepsilon) = \varepsilon_l - \varepsilon_S \qquad (3\text{-}12)$$

بنابراین برای یک نمونه مرتب شده در جهت صعودی خواهیم داشت :

$$Ra(\varepsilon) = \varepsilon_n - \varepsilon_1 \qquad (3\text{-}12a)$$

عرض یک نمونه را می توان از مجموعه تعاریف $D\equiv\{d_j\ ,\ j=1,\dots,m\}$ ، آن نیز بدست آورد .
در اینصورت روابط نظیر روابط 3-12 و 3-12a عبارت خواهد بود از :

$$Ra(\varepsilon)=Ra(D)=d_l-d_S \qquad (3\text{-}13)$$

$$Ra(\varepsilon)=Ra(D)=d_m-d_1 \qquad (2\text{-}13a)$$

مثال 3-9

در مثال 3-8 نمونه مرتب شـده $n=9$ $\varepsilon\equiv(1\ ,1\ ,1\ ,1\ ,1\ ,2\ ,2\ ,2\ ,4)$ را داریـم . مجموعه تعاریف آن $m=3$, $D\equiv\{1, 2\ , 4\}$ می باشد . برای تعیـین عـرض آن از فرمـول 3-12a ویا 3-13a استفاده می کنـیم . 4-

$$Ra(\varepsilon)=\varepsilon_n-\varepsilon_1=1=3$$

$$Ra(\varepsilon)=d_m-d_1=4\text{-}1=3$$

در خاتمه این بخش مشخصات مختلـف نمونـه داده در مثـال 3-3 را کـه در مثـال هـای فـوق محاسبه شده اند بطور خلاصه بیان می کنیم .

$M=1.\overline{6}$, $S^2=0.\overline{8}$, $S=0.94$, $Med(\varepsilon)=1$, $Ra(\varepsilon)=3$

(وجود علامت (-) بالای آخرین رقم از هر عدد در بالا نشان دهنده تکراری بودن آن رقـم مـی باشد).

3-1-6 نمودارها و پلی گن ها (Histograms and Polygons)

از حالا به بعد تعداد عناصر m یک نمونه ε ، اندازه (size) آن نمونه نامیده خواهد شد . عناصر یک نمونه با اندازه یا سایز بزرگ m اغلب به تعدادی کلاس (Classes) و یـا (Categories) تقسیم می شوند . هر کلاس از گروه n_i تا عنصر $(n_i < n)$ تـشکیل شـده اسـت . بـرای کلاسـه بندی یا گروه بندی هر نمونه معمولاً عـرض (Range) آنرا تعـیین می کـنند . ســپس عـرض نمـونه را به تعداد k فاصله بوسـیله k+1تـا مـرز (Class-Boundary) یـا (Class-Limit) تقسیم می کنند . معمولاً فاصله ها را مساوی انتخاب می کنند . اختلاف بین مرز بالائی و پـایینی هر کلاس را عرض کلاس (Class-width) می نامند . تعداد عناصر داخل هر کلاس را حجـم کلاس (Class-Count) یا (Class-Frequency) می نامند . در آمار اعمـال فـوق را کلاسـه بندی نمونه (classification of the sample) می نامند .نمایش گرافیکی یک نمونـه کلاسـه شده را بوسیله جعبه یا مستطیل هائی کـه هـر کـدام نماینـده یـک کـلاس مـی باشـند نمـودار (هیستوگرام) آن نمونه می نامند .

مثال 10-3

نمونه زیر را در نظر می گیریم :

$\varepsilon \equiv$ (17 , 3 , 2 , 8 , 1 , 5 , 2 , 4 , 6 , 15 , 8 , 9 , 2 , 4 , 10 , 9 , 11 , 12 , 4 , 5 , 8 , 6 , 7, 4,5) , n=25

أ-

فاصله از a تا b به چهار شکل زیر می تواند تعریف شود :

1- فاصله باز که با (a,b) نشان می دهند یعنی (x : a<x<b)

2- فاصله بسته که با [a,b] نشان می دهند یعنی (x : a≤x≤b)

3- فاصله بسته-باز که با (a,b] نشان می دهند یعنی (x : a<x≤b)

4- فاصله باز-بسته که با [a,b) نشان می دهند یعنی (x : a≤x<b)

برای ایجاد توافق بین علائم مورد استفاده در اینجا با قراردادهای تئوری مجموعه ها فواصل تعریف درفوق را مثل یک مجموعه در نظر می گیریم . برای اینکه این فواصل از مجموعه نقاط قابل تمیز باشند ، آنها را مجموعه های فشرده (Compact set) می نامیم .

نخست عرض (Range) نمونه ε را با استفاده از فرمول (12-3) محاسبه می کنیم .

$Ra(\varepsilon) = \varepsilon_l - \varepsilon_S =$ 17-1=16

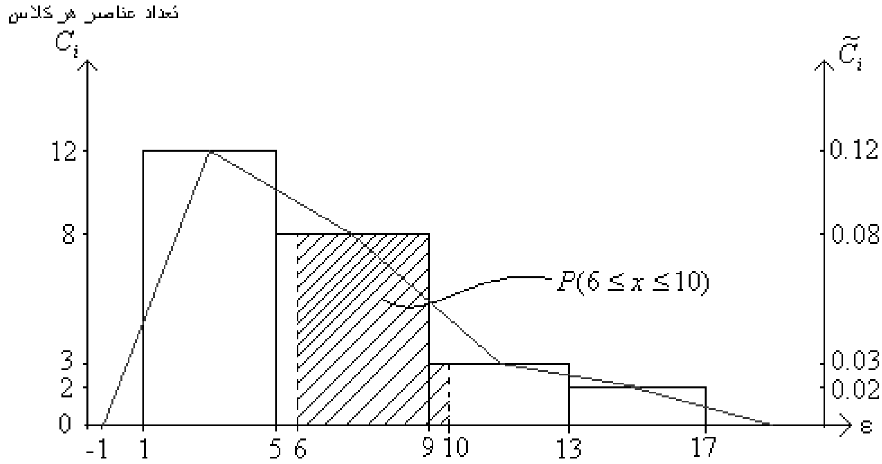
فرض می کنیم که عرض نمونه به چهار فاصله مساوی زیر تقسیم شود

[1 , 5] , (5 , 9] , (9 , 13] , (13 , 17]

تعداد عناصر داخل فواصل فوق عبارتند از :

C_1([1 , 5])=12 , C_2((5 , 9])=8 , C_3((9 , 13])=3 , C_4((13 , 17])=2

نمودار یا هیستوگرام نمونه ε در شکل زیر نشان داده شده است . محور افقی نماینده فاصله ها و مرزهای آنها و محور عمودی نماینده C_i ها می باشند .



شکل ۴-۳

در شکل فوق یک مستطیل روی هر فاصله (کلاس) رسم شده است که ارتفاع آن مساوی تعداد عناصر آن کلاس می باشد . معمولاً خواسته می شود که سطح زیر نمودار (جمع مساحت مستطیل ها)برابر واحد باشد . فرض می کنیم که تعداد k کلاس با عناصر C_i در هر کلاس بطوریکه $\sum_{i=1}^{k} C_i = n$ و عرض (Class-width) آنها مساوی و برابر Δ باشد داریم . در اینصورت مساحت زیر نمودار (a) برابر است با :

$$a = C_1\Delta + C_2\Delta + \cdots + C_k\Delta = \Delta\sum_{i=1}^{k} C_i = n\Delta$$

برای اینکه مساحت زیر نمودار برابر واحد گردد کافیست ارتفاع هر مستطیل را به $n\Delta$ تقسیم کنیم . در اینصورت ارتفاع جدید هر مستطیل برابر $\tilde{C}_i = \frac{C}{n\Delta}$ که آنرا شمارش نسبی می نامند خواهد بود (این شمارش نسبی را با شمارش نسبی در بخش 2-1-3 که احتمال تجربی یک عنصر تنها را نشان می دهد مقایسه کنید. ولی در اینجا ما با شمارش نسبی عناصر موجود در یک فاصله سروکار داریم)

مثال 11-3

در مثال 10-3 داریم n=25 و Δ=4 برای بدست آوردن شمارش های نسبی $\tilde{C}_i$ مقادیر C_i را به nΔ=100 تقسیم می کنیم .

$\tilde{C}_1 = \frac{12}{100} = 0.12$, $\tilde{C}_2 = \frac{8}{100} = 0.08$, $\tilde{C}_3 = \frac{3}{100} = 0.03$,

$\tilde{C}_4 = \frac{2}{100} = 0.02$

با در نظر گرفتن مقادیر جدید ارتفاع مستطیل ها ، شکل نمودار می تواند بـه همـان حـال بـاقی بماند بشرطی که اشل روی محور عـمودی مطابق اندازه های جدید عوض شود . محور قـائم در سمت راست شکـل 4-3 شمارش های جدید را نشان می دهد .

با در نظر گرفتن شمارش های نسبی ، مساحت زیر نمودار (a) برابر واحد می شود

a=4×0.12+ 4×0.08+ 4×0.03+ 4×0.02=1

با محاسبه a که باید برابر واحد گردد صحت مقادیر $\tilde{C}_i$ها ثابت می شود .

فرض کنیم که بزرگترین و کوچکترین کمیت روی محور افقی نمودار را با l و s نشان می دهیم (در شکل 4-3 $l=17$ و $s=1$ می باشد) برای هر فاصـله فرعـی $D'=[a,b]$ از فاصـله $[s,l]$ ، می توان سطح زیر نمودار $\propto(D')$ را محاسبه نمود که برابـر عـددی بـین صـفر و یـک خواهد بود . بدین جهت می توان $\propto$ را تابعی در نظر گرفت که هـر فاصـله فـرعی از $[s,l]$ را روی $[0,1]$ تصویر می کند . بنابراین بسادگی دیده می شود که $\propto$ می توانـد یـک تـابع احتمـال (بخش 1-2) ، یا بطور دقیق تر ،یکی از توابع پخش احتمال PDF نمونه ε باشد . مـسلماً چنـین تابعی ($\propto$) بستگی به طرز کلاسه کردن نمونه خواهد داشت .

از مباحث فوق چنین بر می آید که احتمال هر فاصله فرعی D' از فاصله $[s,l]$ برابر سطح زیـر نمودار روی فاصله فرعی می باشد. از طرف دیگر محـور عمـودی در شـکل 4-3 نـشان دهنـده احتمالات نخواهد بود (دوباره نمودار مزبور را با Bar diagram بخش 2-1-3 مقایسه کنید.)

مثال3-12

با مراجعه به شکل 4-3 ممکن است سؤال شود که احتمال $P(D')$ ، $D'=[6,10]$، یـا احتمـال اینکه یک عنصر نمونه (مثلاً x) بین 6 و 10 باشد چقدر است یا به عبارت دیگر:

$$P(6\le x\le 10) = ?$$

جواب عبارتست از مساحت زیر نمودار در فاصله $[6,10]$ که در شکل 4-3 با هاشور مشخص شده است

$$P(6\le x\le 10)=P(6\le x\le 9)+ P(9<x\le 10)=(9 -6)\times 0.08 +(10-9)\times 0.03 =0.27$$

از طرف دیگر با بررسی روی نمونه واقعی ε در مثال 1-3 معلـوم مـی شـود کـه تعـداد واقعـی عناصر موجود در فاصله $[6,10]$ مساوی 9 بوده که برابـر $\frac{9}{25}\times 100=0.36$ یعنـی 36 درصـد نمونه می باشد . بنابراین احتمال واقعی برابر است $P(6\le x\le 10)= 0.36$ که بـا مقـدار بدسـت آمده از نمودار (0.27) مغایرت دارد .

این اختلاف بیشتر از طرز کلاسه بندی نمونه ناشی می شود . این اختلاف کوچکتر می شود اگـر نمونه به کلاس هائی که مرز آنها ، هیچکدام از عناصر تلاقی نکند، تقسیم شود.

ایجاد یک نمودار روش خاصی ندارد و بستگی به نمونه دارد . در این مقوله بیشتر از این صحبت نخواهد شد .

مثال 3-13

اگر ما کلاسهای زیر را برای نمونه داده شده در مثال 3-10 انتخاب کنیم

$[0.7 , 4.8] \quad , \quad (4.8 , 8.9] \quad , \quad (8.9 , 13] \quad , \quad (13 , 17.1]$

دوباره 4 فاصله مساوی و با عرض $\Delta=4.1$ خواهیم داشت .

شمارش کلاس ها (تعداد عناصر داخل کلاس) عبارتند از :

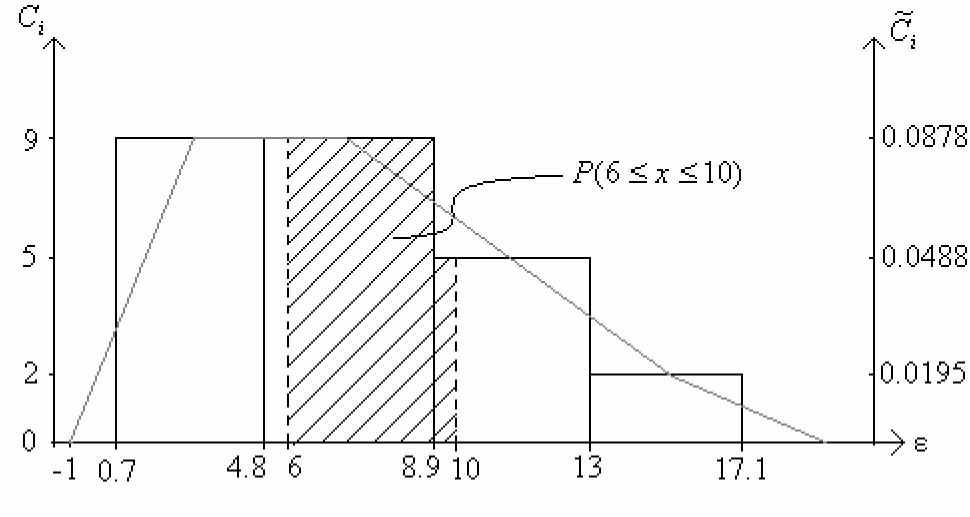
$C_1=9 \quad , \quad C_2=9 \quad , \quad C_3=5 \quad , \quad C_4=2$

و کمیت $n\Delta$ برابر $n\Delta=25\times4.1=102.5$ خواهد بود . شمارش های نسبی عبارتند از :

$$\tilde{C}_1=\tilde{C}_2=\frac{9}{102.5}=0.0878 \quad , \quad \tilde{C}_3=\frac{5}{102.5}=0.0488 \quad ,$$

$$\tilde{C}_4=\frac{2}{102.5}=0.0195$$

نمودار جدید نمونه ε در شکل 3-5 دیده می شود .



شکل ۵-۳

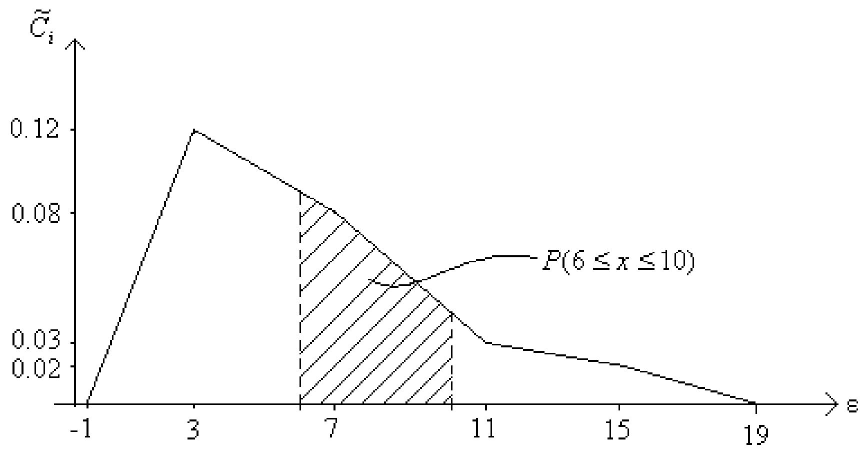
احتمال $P(6\le x\le 10)$ ، سطح هاشور زده شده چنین محاسبه می شود

$$P(6\le x\le 10)=P(6\le x\le 8.9)+P(8.9<x\le 10)=2.9\times0.0878+1.1\times0.0488=0.31$$

مقدار فوق خیلی نزدیکتر به احتمال واقعی (0.36) ، در مقایسه با عدد 0.27 که از نمودار قدیمی بدست آمده است می باشد .

نمایش یک نمودار را بوسیله نقاط مرکزی هر فاصله (Class-Midpoint) و ارتفاع آن نقاط (شمارش نسبی مربوط به هر کلاس) یک پلی گن مینامند . برای اینکه سطح زیر پلی گن

مساوی واحد باشد مجبوریم دو فاصله یکی در ابتدا و دیگری در انتهای نمودار اضافه کنیم . بوسیله نقاط مرکزی این دو فاصله، s',l' ، پلی گن بسته می شود. بسادگی دیده می شود



شکل ۶-۳

که سطح زیر پلی گن ($\propto'$) نیز دارای خصوصیات تابع احتمال می باشد. یعنی $\propto'$ یکی از توابع پخش احتمال PDF نمونه می باشد . بنابراین $\propto'$ می تواند برای تعیین احتمال هر فاصله فرعی $D' = [a,b] \subset [s',l']$ مورد استفاده قرار گیرد . باید در نظر داشت که در پلی گن نیز ارتفاعات نماینده احتمالات نیستند .

مثال 3-14

پلی گن مربوط به نمودار شکل 4-3، در شکل 6-3 نمایش داده شده است .

نظیر نمودار ، سطح زیر پلی گن یعنی a باید مساوی واحد باشد . این شرط در پلی گن شکل 6-3 برقرار است

$$a = 4\left(\frac{1}{2}\times 0.12 + \frac{1}{2}(0.12+0.08) + \frac{1}{2}(0.08+0.03) + \frac{1}{2}(0.03+0.02) + \frac{1}{2}\times 0.02\right).$$
$$= 2(0.12+0.20+0.11+0.05+0.02) = 1.00$$

برای تعیین احتمال $P(6 \le x \le 10)$ مساحت زیر پلی گن را در فاصله [6,10] ، سطح هاشور زده شده در شکل 6-3 ، محاسبه می کنیم . برای اینکار اول ارتفاع پلی گن را در نقاط 6 و 10 محاسبه می کنیم که به ترتیب برابرند با 0.090 و 0.0425 ، سپس داریم

$$P(6 \le x \le 10) = P(6 \le x \le 7) + P(7 < x \le 10) = 1\times\frac{1}{2}(0.09+0.08) + 3\times\frac{1}{2}(0.08+0.0425)$$
$$= 1\times 0.085 + 3\times 0.06125 = 0.27$$

نتیجه فوق برابر است با مقداری که توسط نمودار برای احتمال فوق بدست آمده است .

تا اینجا نمودار و پلی گن مربوط به احتمال یک نمونه را ایجاد کردیم . نظیر همین کارها را می توان در مورد نمودار و پلی گن مربوط به تابع CDF نمونه که آنها را به ترتیب Cumulative Histogram و Cumulative polygon می نامیم انجام داد . در اینجا از تابع CDF ، فرمول 3-2 ، که شکل زیر تغییر یافته است استفاده می کنیم . (یادآوری فرمول 3-2 : $(C(d_i) = \sum_{j \le i} P(d_j) \in [0,1]$

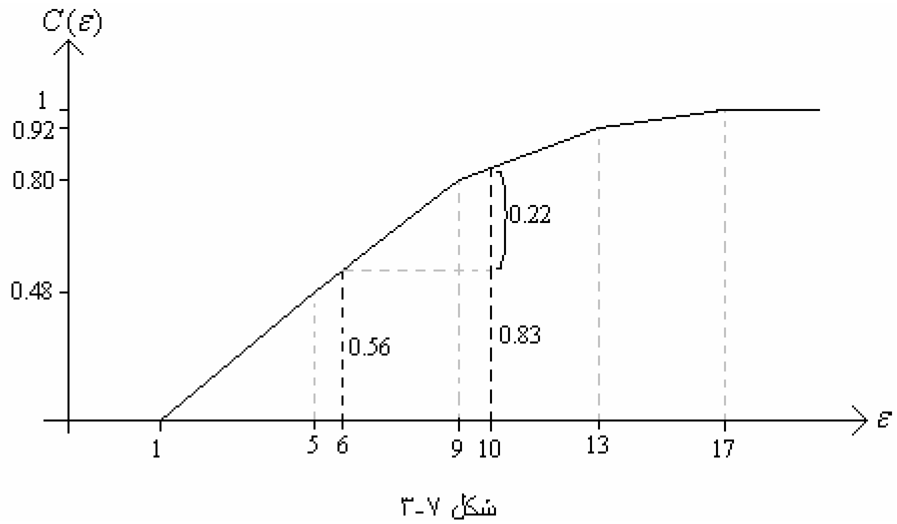
$$C(a) = P(x \le a) = \sum_{x_i \le a} P(x_{i-1} < x < x_i) \qquad (3\text{-}14)$$

مثال 3-15

در این مثال نمودار تجمعی و پلی گن تجمعی نمونه داده شده در مثالهای فوق را رسم می کنیم برای رسم نمودار تجمعی از شکل 3-4 استفاده می کنیم (بخاطر داشته باشید که در استفاده از نمودار ، احتمال هر یک از عناصر به تنهایی همیشه مساوی صفر است)

C(1) = P(1)=0
C(5) = P[1 , 5] = 4×0.12 = 0.48
C(9) = C(5)+P(5 , 9] = 0.48 + 4×0.08 = 0.48 + 0.32 = 0.80
C(13) = C(9)+P(9 , 13] = 0.80 + 4×0.03 = 0.80 + 0.12= 0.92
C(17) = C(13)+P(13 , 17] = 0.92 + 4×0.02 = 0.92 + 0.08= 1

شکل 3-7 نمایش نمودار تجمعی فوق می باشد .



شکل ۷-۳

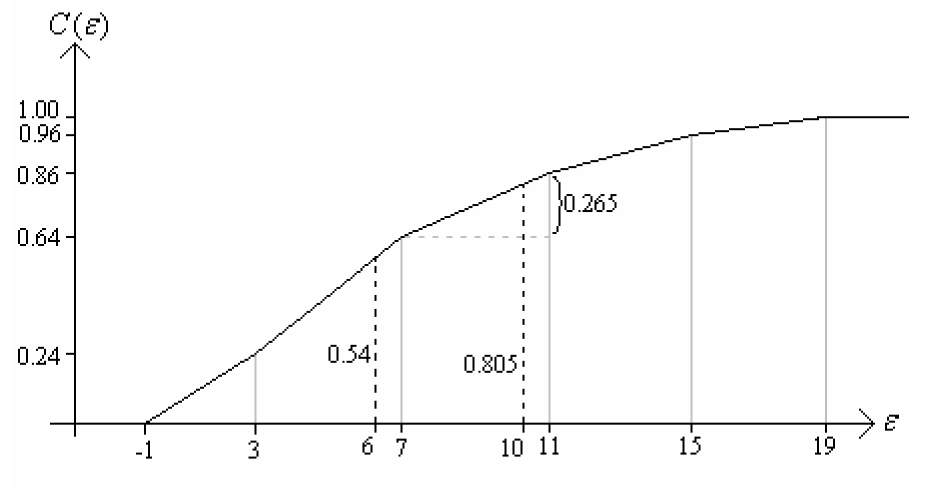
برای رسم پلی گن تجمعی محاسبات زیر را با استفاده از شکل 3-6 انجام می دهیم

C(-1) = 0

C(3) = P[-1 , 3] = $\frac{1}{2}(4 \times 0.12) = 0.24$

$$C(7) = C(3) + P(3,7] = 0.24 + \frac{1}{2} \times 4(0.12 + 0.08) = 0.64$$
$$C(11) = C(7) + P(7,11] = 0.64 + \frac{1}{2} \times 4(0.08 + 0.03) = 0.86$$
$$C(15) = C(11) + P(11,15] = 0.86 + \frac{1}{2} \times 4(0.03 + 0.02) = 0.96$$
$$C(19) = C(15) + P(15,19] = 0.96 + \frac{1}{2} \times 4 \times 0.02 = 1.00$$

شکل 8-3 نمایش پلی گن فوق می باشد . در این شکل و در شکل 7-3 خصوصیات یک تابع CDF (مثال 4-3) نمایان است .



شکل ۸-۳

با توجه به اشکال 7-3 و 8-3 ملاحظه می شود که پلی گن تجمعی از نقاط مرکزی هر کلاس و ارتفاع آنها در نمودار تجمعی استفاده می کند ، بنابراین ارتباط بین نمودار تجمعی و پلی گن تجمعی آن نظیر ارتباط بین نمودار و پلی گن می باشد .

با توجه به طبیعت تابع CDF ملاحظه می شود که یک احتمال کل ، سطح زیر منحنی PDF مشخص می شود از نقطه شروع تا یک نقطه مورد نظر ، فقط با ارتفاع آن نقطه در منحنی CDF مشخص می شود . بنابراین نمودار و یا پلی گن تجمعی را می توان برای تعیین احتمال P[a,b] ، a<b ، مورد استفاده قرار داد. این احتمال برابر است با ارتفاع نقطه b ، در منحنی ، در منحنی CDF ، منهای ارتفاع نقطه a .

مثال 3-16

مقدار احتمال P[6,10] را با استفاده از

أ- نمودار تجمعی در شکل 7-3

ب-پلی گن تجمعی در شکل 8-3

محاسبه کنید .

أ- ارتفاعات مربوط به نقاط 6 و10 در شکل 7-3 ، از راه تناسب ، به ترتیب عبارتند از 0.56 و 0.83 بنابراین داریم :

$$P[6,10] = P(6 \le x \le 10) = 0.83 - 0.56 = 0.27$$

این مقدار همانست که از پلی گن مثال 14-3 بدست آمده است

ب- ارتفاعات مربوط به 6 و10 در شکل 8-3 ۀ از راه تناسب ، به ترتیب عبارتند از 0.54 و 0.805 بنابراین داریم :

$$P[6,10] = P(6 \le x \le 10) = 0.805 - 0.54 = 0.27$$

مقدار بدست آمده دوباره همان است که از پلی گن مثال 14-3 بدست آمده است .

در خاتمه این مبحث ، باید گفت که با انتخاب کلاس های با عرض کوچکتر و مناسب تر برای یک نمونه ، می توان هر دو نمودار و پلی گن (غیر تجمعی یا تجمعی) تصفیه شده بدست آورد . در عمل تصفیه کردن ، اشکال صاف تر دیده می شوند .

3-2 آمار یک متغیر تصادفی (Random Variable)

3-2-1 تابع تصادفی و متغییر تصادفی

به منظور ایجاد سهولت و سرعت بیشتر در حل مسائل مربوط به تعیین احتمال فاصله (رجوع شود به نمودار و پلی گن بخش 6-1-3) علم آمار روش مناسبتری را توسعه داده است . در این روش بجای تابع پخش احتمال عددی ، که روی مجموعه تعاریف گسسته از یک نمونه تعریف می شوند ، توابع مناسبتری بکار گرفته می شوند . برای اینکار ما نخست به تعریف دو ایده ال : تابع تصادفی و متغیر تصادفی ، از دنیای واقعی می پردازیم .

یک تابع تصادفی عبارتست از تابع x که یک مجموعه نا مشخص * U را در مجموعه R تصویر می کند .

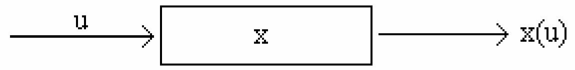
$$x \in \{U \to R\}$$

(بعداً در مبحث چند متغیری تابع $x \in \{U \to R^m\}$ که R^m عبارتست از توان m ام مجموعه R بسط داده می شود).

==

* در علوم تجربی U ممکن است بطور کامل و یا قسمتی از آن معلوم باشد. علوم آماری فرض می کند که مجموعه U ناشناخته است و یا با قسمتی ناشناخته از آن کار می کند

رابطه فوق اینطور تعبیر می شود که به ازای هـر $u \in U$ ، تـابع x دارای مقـدار $x(u) \in R$ مـی باشد. اما به علت نا مشخص بودن مجموعه U نمی توان فرمول مشخصی برای x پیدا کرد و مـا برای نشان دادن اینکه مبحث تابع تصادفی خالی از استفاده نیست بشرح تجربه زیر می پردازیم. فرض کنید که تابع x عبارتست از دستگاهی یا سیستمی که هر بار بکار مـی افتـد یـک مقـدار x(u) تولید می کند.



بدون اطلاع از ساختمان داخلی دستگاه تنها کاری که مـی تـوان انجـام داد ثبـت مقـادیر x(u) بیرون آمده از دستگاه می باشد . وقتی به تعداد زیاد و کـافی از مقـادیر x(u) ثبـت شـد ، مـی توانیم نموداری که شمارش نسبی x(u) را در هر فاصله $[x_0 , x_1]$ نشان می دهد رسـم کنـیم . برای اینکار فرض باندازه ای زیاد مقادیر x(u) جمع آوری کرده ایـم کـه مـی تـوانیم شـمارش نسبی هر فاصله اختیاری کوچک dx را محاسبه کرده در نتیجـه نمـودار یـا هیـستوگرام صـاف بدست بیاوریم . حد شمارش نسبی در فاصله [x , x+dx] تقسیم بر عرض فاصله ، dx ، وقتـی dx بسمت صفر میل می کند . کمیتی است تابع x که آنرا با $\Phi(x)$ نشان می دهیم . ایـن تـابع $x \in R$ را در R تصویر می کند . بنابراین به ازای یک مقـدار از x(u) یـک مقـدار بـرای تـابع $\Phi(x)$ وجود خواهد داشت . با زبان ریاضی می شود گفت که ماحصل دستگاه فوق هـر بـار کـه روشن می شود یک زوج $(x(u), \Phi(x))$ می باشد که آنرا متغیر تصادفی می نامند . مرسوم است که x(u) را به تنهائی ، با در نظر گرفتن اینکه تابع $\Phi(x)$ مربوطـه نیـز معلـوم اسـت ، متغیـر تصادفی می نامند .

ملاحظه می شود که تابع Φ دارای مقادیری مثبت و یا صفـر کـه متعلـق بـه مجموعـه اعـداد حقیقی R می باشند اتخاذ می کند . بعبارت دیگر تابع $\Phi(x)$ غیر منفی و روی R تعریـف شـده است . بعلاوه خودمان را منحصر به نوعی از توابع Φ خواهیم کرد که قابل انتگـرال گیـری روی R باشند یعنی دست کم روی R پیوسته باشند .

3-2-2 تابع پخش احتمال PDF و CDF یک متغیر تصادفی

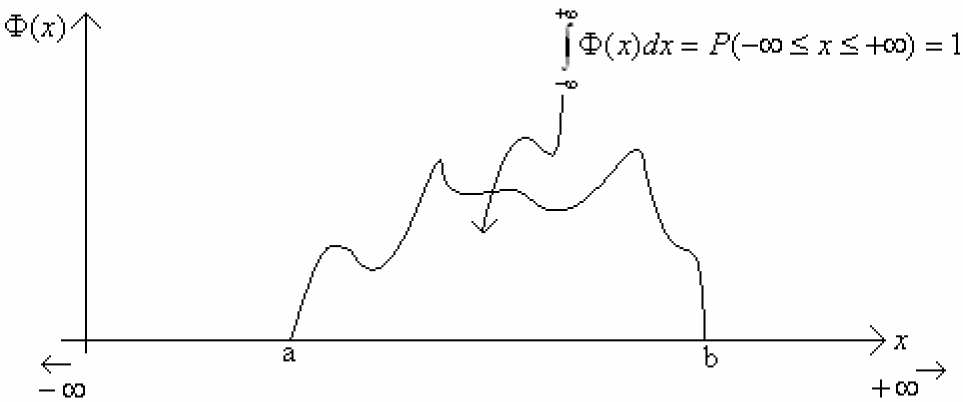
تابع Φ متعلق به متغیر تصادفی x را که در بخش 1-2-3 مورد بحث بود تابع پخش احتمال x مینامند. می توان آنرا معادل تابع پخش احتمال تجربی یک نمونه دانست . از تعریـف تـابع Φ در بخش گذشته معلوم می شود که :

.

$$\int_{-\infty}^{+\infty} \varphi(x)dx = 1 \qquad (3\text{-}15)$$

چون سطح زیر نمودار صاف شده باید مساوی واحد باشد .

این خاصیت سوم تابع PDF ، Φ، می باشد جائیکه قابل انتگرال گیری و غیر منفی بودن خواص اول دوم می باشند . ملاحظه می کنیم که معادله 3-15 عبارتست از شرط لازم برای اینکه $\Phi(x)dx$ یک احتمال نامیده شود . در شکل زیر یک PDF نشان داده شده است . مقدار انتگرال 3-15، سطح هاشور زده شده مشخص شده است .



شکل ۹-۳

انتگرال معین تابع Φ روی فاصله $D' \subset D$ را احتمال $P(D')$ می نامند . در زیر احتمال چند فاصله داده شده است .

$$\int_{-\infty}^{x_0} \Phi(x)dx = P(x \le x_0) \in [0,1] \qquad (3\text{-}16a)$$

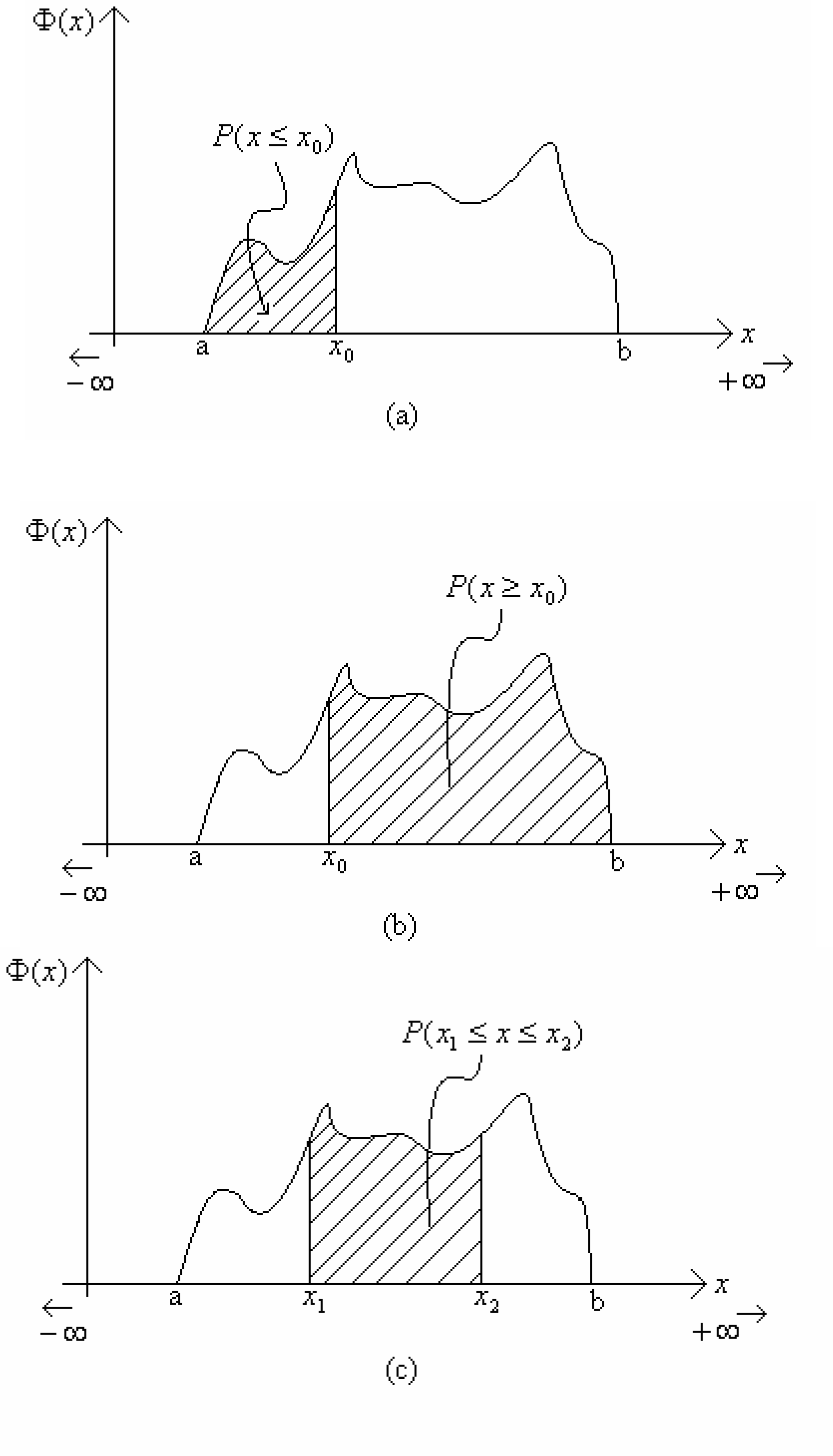
$$\int_{x_0}^{\infty} \Phi(x)dx = P(x \ge x_0) \in [0,1] \qquad (3\text{-}16b)$$

$$\int_{x_1}^{x_2} \Phi(x)dx = P(x_1 \le x \le x_2) \in [0,1] \qquad (3\text{-}16c)$$

در نتیجه خواهیم داشت :

$$P(x \ge x_0) = 1 - P(x \le x_0) \qquad (3\text{-}17)$$

انتگرالهای 3-16a و 3-16b و3-16c در اشکال زیر با هاشور مشخص شده اند .



شکل ۱۰-۳

در این مرحله فرق بین فضای احتمال گسسته و فضای احتمال پیوسته باید در نظرمان مجسم شود . در فضای احتمال گسسته مقدار تابع PDF در هر نقطه از مجموعه گسسته تعاریف D همان احتمال آن نقطه می باشد . در صورتیکه در فضای پیوسته سطح زیر PDF است که مقدار

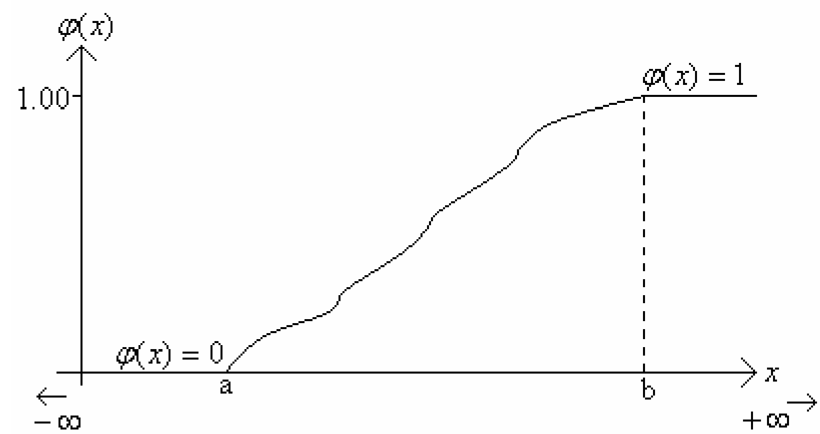
احتمال را نشان می دهد . ما قبلاً با این مسئله در نمودارها آشنا شدیم . برای توضیح بیشتر به رابطه زیر توجه کنید :

$$P(x = x_0) = \int_{x_0}^{x_0} \Phi(x)dx = 0$$

شبیه بخش 2-1-3 تابع پخش احتمال تجمعی φ را بصورت زیر تعریف می کنیم :

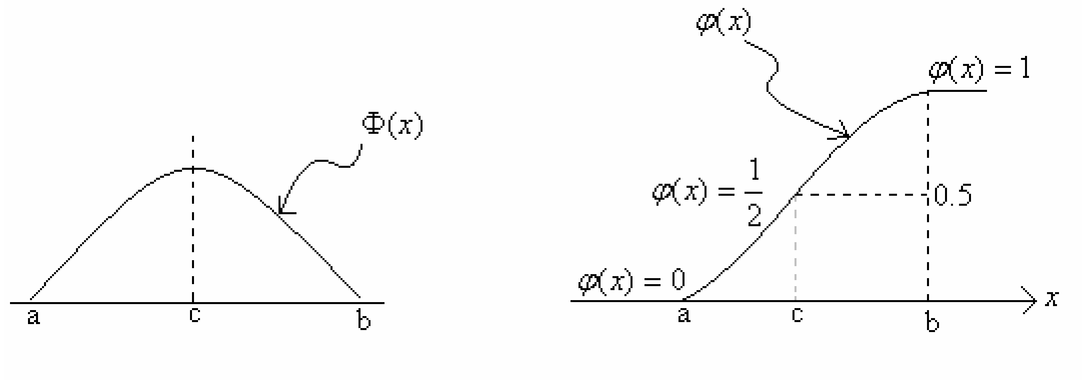
$$\varphi(x) = \int_{-\infty}^{x} \Phi(y)dy \in [0,1] \qquad (3\text{-}18)$$

در اینجا y متغیر داخل انتگرال است که از $-\infty$ تا x تغییر می کند . متغیر داخل انتگرال را برای اینکه با حد بالای انتگرال یعنی x اشتباه نشود y قرار داده ایم . شکل 11-3 فرم تابع φ ، CDF ، متعلق به PDF رسم شده در شکل 9-3 را نشان می دهد.



شکل ۱۱-۳

اگر $\Phi(x)$ یک تابع قرینه باشد ، در آن صورت تابع $\varphi(x)$ نسبت به محور $\varphi(x)=\frac{1}{2}$ قرینه معکوس خواهد بود . شکل 12-3 مثالی در این زمینه را نشان می دهد .



شکل ۱۲-۳

ملاحظه می شود که φ تابع اولیه Φ می باشد

$$\Phi(x) = \frac{d\varphi(x)}{dx}$$

بعلاوه برای برقراری شرط

$$\int_{-\infty}^{+\infty} \Phi(x)dx = 1$$

تابع $\Phi(x)$ باید به ازای مقادیر بینهایت بزرگ و بینهایت کوچک x بسمت صفر میل کند.

بنابراین خواهیم داشت :

$$\lim_{x \to +\infty} \varphi(x) = 1$$

$$\lim_{x \to -\infty} \varphi(x) = 0 \quad ,$$

3-2-3 میانگین و واریانس یک متغیر تصادفی

ملاحظه می شود که اگر ما تابع پخش احتمال تصادفی را ندانیم و یا تابع پخش برای آن فرض نکنیم ، اصولاً مبحث متغیر تصادفی بی فایده خواهد بود . از طرف دیگر ارتباط یک به یک بین متغیر تصادفی و تابع پخش آن ، بطوریکه برای یک نمونه و تابع پخش آن موجود بود وجود ندارد . متغیر تصادفی تنها بعنوان متغیر تابع پخش خود عمل می کند .

متغیر تصادفی که مقدارآن از $-\infty$ تا $+\infty$ تغییر می کند می تواند بعنوان متغیر تابع PDF در نظر گرفته شود . بنابراین نمی توانیم دقیقاً با همان مفهومی که در مورد میانگین و واریانس یک نمونه صحبت می کردیم در اینجا نیز صحبت کنیم . اما در عوض می توانیم از مرکز ثقل سطح زیر منحنی PDF بعنوان میانگین صحبت کنیم و همینطور در مورد واریانس نیز می توانیم صحبت کنیم .

میانگین μ یک متغیر تصادفی x طبق رابطه زیر تعریف می شود:

$$\mu = \int_{-\infty}^{+\infty} x\Phi(x)dx \tag{3-20}$$

رابطه فوق شبیه رابطه 3-4 در بخش 3-1-3 می باشد. μ اغلب بوسیله اپراتور E^* نشان داده می شود

$$E^*(x) = \mu = \int_{-\infty}^{+\infty} x\Phi(x)dx \quad * \tag{3-21}$$

اپراتور E^* نظیر اپراتور E دارای خواص زیر میباشد:

1) $E^*(kx)= k\, E^*(x)$

2) $E^*(\sum_{j=1}^{r} x^j) = \sum_{j=1}^{r} E^*(x^j)$

در اینجا j=1,…,3 و x^j عبارتند از تعداد دو متغیر تصادفی مختلف که هر کدام دارای تابع پخش احتمال PDF خود میباشند.

همینطور میتوان رابطه زیر را نیز در مورد E^* تعریف کرد:

3) $E^*(E^*(x)) = E^*(x) = \mu$

واریانس σ^2 متغیر تصادفی x با میانگین μ طبق رابطه زیر تعریف میشود:

$$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 \Phi(x) dx \qquad (3\text{-}22)$$

* E^* مثل مخفف Mathematical Expactation یا امید ریاضی است. علامت ستاره برای تمیز دادن دو نوع جمع ∑و∫ می باشد.E عبارت است از امید ریاضی وقتی ∑ مورد استفاده باشد و E^* عبارت است از امید ریاضی وقتی ∫ مورد استفاده قرار گیرد.

رابطه 22-3 شبیه رابطه8- 3 میباشد. جذر واریانس نسبی σ را انحراف معیار (Standard deviation) متغیر تصادفی x می نامند.

رابطه22-3 را بصورت زیر بسط میدهیم :

$$\sigma^2 = \int_{-\infty}^{+\infty} [x^2\Phi(x) - 2x\mu\Phi(x) + \mu^2\Phi(x)]dx$$

$$= \int_{-\infty}^{+\infty} x^2\Phi(x)dx - 2\mu \int_{-\infty}^{+\infty} x\Phi(x)dx + \mu^2 \int_{-\infty}^{+\infty} \Phi(x)dx$$

دیده می شود که دومین انتگرال در سمت راست برابر μ و سومین انتگرال برابر واحد میباشد با جایگزینی مقادیر انها خواهیم داشت :

$$\sigma^2 = \int_{-\infty}^{+\infty} x^2\Phi(x)dx - \mu^2 \qquad (3\text{-}23)$$

این رابطه نظیر رابطه 9-3 میباشد. با در نظر گرفتن تعریف اپراتور E^* ، رابطه 22-3 به صورت زیر در می آید:

$$\sigma^2 = E^*\left((x-\mu)^2\right) = E^*\left((x - E^*(x))^2\right) \qquad (3\text{-}23a)$$

واریانس ، σ^2 ، را معمولا به شکل فوق نشان می دهند.

ممان r ام تابع PDF (یا متغیر تصادفی) را به صورت زیر تعریف می کنند

$$m_r = \int_{-\infty}^{+\infty} x^r \Phi(x) dx \qquad (3\text{-}24)$$

به طور دقیق m_r را ممان r ام تابع φ(x) حول صفر می نامند.
از طرفی ممان مرکزی r ام تابع PDF به صورت زیر داده می شود:

$$m'_r = \int_{-\infty}^{+\infty} (x-\mu)^r \Phi(x) dx \qquad (3\text{-}25)$$

با مقایسه تعاریف m_r و m'_r با روابط (3-20) و (3-22) می بینیم که

$$\mu = m_1 \qquad (3\text{-}26a)$$

$$\sigma^2 = m'_2 = m_2 - \mu^2 = m_2 - m_1{}^2 \qquad (3\text{-}26b)$$

نتایج فوق ، (3-26b,a)، را با مشابه های مکانیکی یاد شده در بخش 3-31 مقایسه کنید.

3-2-4 فرض اساسی آمار ، امتحان کردن(Test)

در آمار فرض بر اینست که هر نمونه تـصادفی زائیـده یـک جامعـه(population) اسـت.هـر جامعه با یک متغیر تصادفی XER با دامنه تغیرات بینهایت مشخص شده وانرا متغیـر تـصادفی مادر مینامند .

در آمار مرسوم است که تابع PDF محاسبه مربوطه را برای نمونه تصادفی در نظر گرفته آنـرا تابع PDF مفروض مینامند. ممکن است ارزش آماری چنین فرضی مورد امتحان قرار گیرد.

برای اینکه بتوان از نظر آمار ارزش فرضیه را تعیین باید نمونه ای داشت کـه از جامعـه گرفتـه شده و در آن هر عنصر به طور مستقل از بقیه عناصر انتخاب شده باشد .

این خاصیت اخیر طبق تعریف استاندارد نمونه در آمـار بایـد در یـک نمونـه موجـود باشـد. از آنجاییکه ما فعلا به کار تست می پردازیم بنابراین به همان تعریف اولیه نمونه اکتفا میکنیم.

بینهایت خانواده از توابع PDF وجود دارد هر خانواده از این توابع بوسیله یک و یا چند پـارامتر، که مقدار آنها شکل تابع PDF را مشخص میکنند، تعریف میشود.

شکل هر تابع از یک خانواده نسبت به شکل تابع دیگر از آن خانواده بر حسب مقادیر پارامترها فرق میکند.

معمولا ، در صورت امکان ، از میانگین و انحراف معیـار بـه عنـوان پارامترهـای PDF اسـتفاده میکنند. هر چه تعداد پارامترهای یک خانواده از توابع کمتر باشد کار با آن توابع سهل تر است.

روش معمول اینست که ما اول یک خانواده مناسب از توابعPDF بر اساس تجربه ای که نمونـه تصادفی متکی بر آنست برای نمونه مزبور انتخاب میکنیم. سپس مقادیر پارامترهای آنـرا چنـان

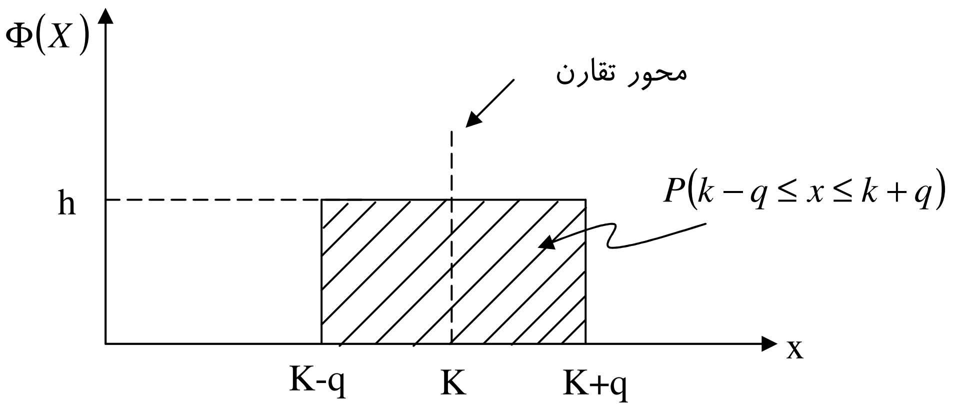
تعیین می کنیم که به بهترین وجهی به نمونه تصادفی منطبق گردد . به عبارت دیگر شکل تابع $\varphi(x)$ اول انتخاب می شود سپس مقدار پارامترهای مربوطه بوسیله بعضی روشهای شناخته شده محاسبه میشوند.

از آنجاییکه منبعد با نمونه های تصادفی توأم با متغیرهای تصادفی مربوطه (جامعه) سر و کار خواهیم داشت ، حروف لاتین را برای نشان دادن مشخصات نمونه و حروف یونانی مترادف لاتینی را برای نشان دادن مشخصات متغیر تصادفی مورد استفاده قرار خواهیم داد. البته تا کنون نیز از این قاعده پیروی کرده ایم.

3-2-5 دو مثال از یک متغیر تصادفی

مثال 7-13

به عنوان اولین مثال فرض می کنیم که یک متغیر تصادفی x دارای تابع PDF مستطیل شکل (یکنواخت) که نسبت به x=k قرینه است ، میباشد در ضمن فرض می کنیم که احتمال های p(x<k-q) و p(x>k+q) مساوی صفر باشند . مسلما تابع PDF فوق دارای فرم زیر خواهد بود:



(شکل 13-3)

$$\Phi(x) = \begin{cases} h & k-q \le x \le k+q \quad for \\ 0 & x < k-q, x > k+q \quad for \end{cases}$$

و یا به طور اختصار:

$$\Phi(x) = \begin{cases} h & |x-k| \le q \\ 0 & |x-k| \le q \end{cases}$$

تابع $\varphi(x)$ فوق ظاهرا دارای سه پارامتر k,q,h میباشد ولی دو تای انها مستقل بوده و آن یکی را می توان با استفاده از شرط 3-15 حذف نمود، یعنی شرط

$$\int_{-\infty}^{+\infty}\Phi(x)dx=1$$

باید برای هر تابع $\varphi(x)$ به عنوان PDF بر قرار باشد، فرض کنیم که پارامتر h را بدین ترتیب حذف کنیم:

$$\int_{-\infty}^{+\infty}\Phi(x)dx=\int_{-\infty}^{k-q}\Phi(x)dx+\int_{k-q}^{k+q}\Phi(x)dx+\int_{k+q}^{\infty}\Phi(x)dx$$

$$=0+\int_{k-q}^{k+q}\Phi(x)dx+0=h\int_{k-q}^{k+q}dx=h[x]_{k-q}^{k+q}=2hq=1$$

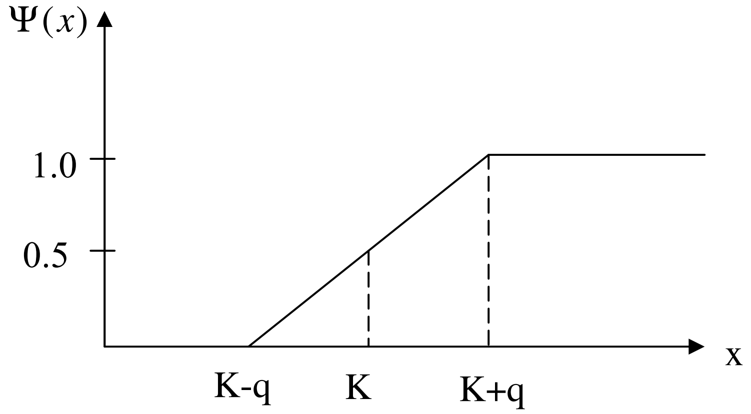
در نتیجه داریم h=1/2q و تابع $\varphi(x)$ به صورت زیر در می آید

$$\Phi(x)=\begin{cases}1/2q & |x-k|\le q \quad for\\ 0 & |x-k|>q \quad for\end{cases}$$

تابع تجمعی CDF مربوط به PDF عبارت خواهد بود از

$$\Psi(x)=\begin{cases}0 & x\le k-q \quad for\\ 1/2q\int_{k-q}^{x}dx=1/2q(x-k+q) & |x-k|\le q \quad for\\ 1 & x>k+q \quad for\end{cases}$$

تابع فوق در شکل 3-14 نمایش داده شده است



(شکل 3-14)

در شکل فوق دیده میشود که تابع $\psi(x)$ در فاصله ای که $\Phi(x)\neq 0$ خطی بوده و در سایر جاها ثابت میباشد و ملاحظه میشود که

$$\Phi(x)=\frac{d\Psi(x)}{dx}$$

میانگین PDF از فرمول 20-3 محاسبه میشود.

$$\mu=\int_{-\infty}^{+\infty} x\Phi(x)dx=1/2q\int_{k-q}^{k+q} xdx=1/2q\left[x^2/2\right]_{k-q}^{k=q}$$

$$=1/4q(k^2+2kq+q^2-k^2+2kq-q^2)=4kq/4q=k$$

نتیجه فوق پیش بینی ما را که از تابع PDF قرینه نسبت به محور x=k داشتیم تأیید میکند.

واریانس تابع PDF از فرمول 22-3 محاسبه می شود

$$\sigma^2=\int_{-\infty}^{+\infty}(x-\mu)^2\Phi(x)dx=1/2q\int_{k-q}^{k+q}(x-k)^2dx=1/2q\int_{k-q}^{k+q}x^2dx+k^2/2q\int_{k-q}^{k+q}dx$$

$$-k/q\int_{k-q}^{k=q}xdx=1/2q\left[x^3/3\right]_{k-q}^{k+q}-2k^2+k^2$$

$$=1/6q(k^3+3k^2q+3kq^2+q^3-k^3+3k^2q-3kq^2+q^3)-k^2$$

$$=6k^2q/6q+2q^3/6q-k^2=q^2/3$$

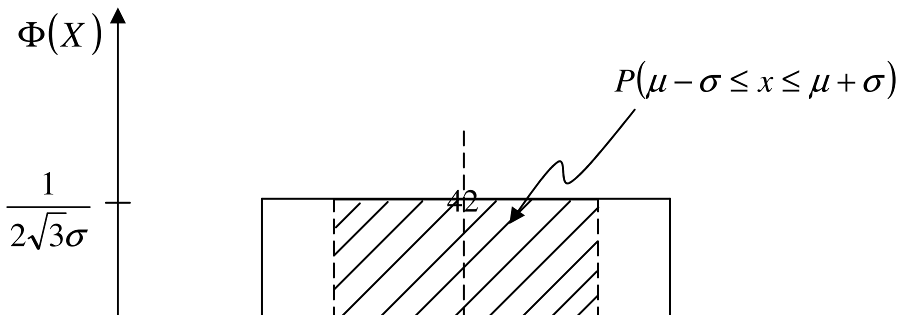
از آنجائیکه داریم $k=\mu$ و $q=\sqrt{3}\sigma$ و $h=\frac{1}{2q}=\frac{1}{2\sqrt{3}\sigma}$ ، می توانیم تابع مستطیل شکل را که با R نشان می دهیم برحسب میانگین و واریانس آن بنویسیم.

$$R(\mu,\sigma;x)=\Phi(x)=\begin{cases}\frac{1}{2\sqrt{3}\sigma} & \text{برای } |x-\mu|\le\sqrt{3}\sigma\\ 0 & \text{برای } |x-\mu|>\sqrt{3}\sigma\end{cases}$$

همینطور می توانیم تابع CDF مربوطه را که با R_C نشان می دهیم برحسب σ,μ بنویسیم

$$R_C(\mu,\sigma;x)=\psi(x)=\begin{cases}0 & \text{برای } x\le\mu-\sqrt{3}\sigma\\ \frac{1}{2\sqrt{3}\sigma}\left(x-\mu+\sqrt{3}\sigma\right) & \text{برای } |x-\mu|\le\sqrt{3}\sigma\\ 1 & \text{برای } x\ge\mu+\sqrt{3}\sigma\end{cases}$$

فرض می کنیم که می خواهیم احتمال $x\in[\mu-\sigma,\mu+\sigma]$ یعنی $P(\mu-\sigma\le x\le\mu+\sigma)$ را محاسبه کنیم که در آن x دارای پخش احتمال مستطیل شکل می باشد. برای اینکار از فرمول c 16-3 و شکل 15-3 استفاده می کنیم.



$$P(\mu-\sigma \le x \le \mu+\sigma)=\int_{\mu-\sigma}^{\mu+\sigma}\Phi(x)dx=\int_{\mu-\sigma}^{\mu+\sigma}\frac{1}{2\sqrt{3}\sigma}dx=\frac{1}{2\sqrt{3}\sigma}[2\sigma]$$

$$=\frac{1}{\sqrt{3}}=\frac{\sqrt{3}}{3}=\frac{1.732}{3}\cong 0.577\cong 0.58$$

احتمال فوق برابر سطح هاشور زده در شکل 15-3 می باشد.

همینطور برای تابع PDF یکنواخت فوق خواهیم داشت

$$P(\mu-2\sigma \le x \le \mu+2\sigma)\equiv P(\mu-3\sigma \le x \le \mu+3\sigma)\equiv 1.0$$

در تست آماری ما اغلب احتیاج داریم که ممانهای PDF (بخش 3-2-3) را محاسبه کنیم.

برای مثال ممان سوم m_3 حول صفر تابع مستطیل شکل فوق چنین محاسبه می شود.

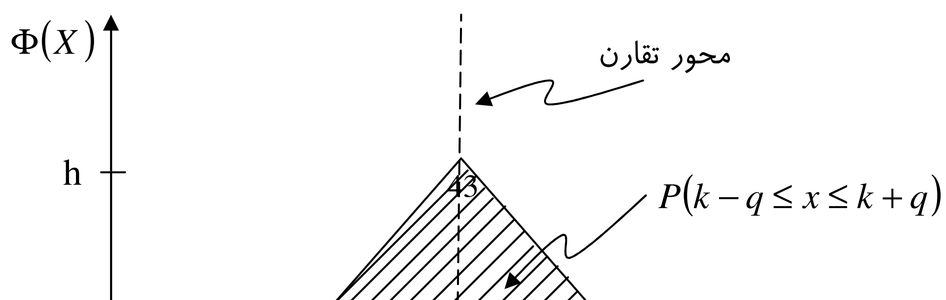
$$m_3=\int_{-\infty}^{+\infty}x^3\Phi(x)dx=\int_{\mu-\sqrt{3}\sigma}^{\mu+\sqrt{3}\sigma}x^3\frac{1}{2\sqrt{3}\sigma}dx=\frac{1}{2\sqrt{3}\sigma}\left[\frac{x^4}{4}\right]_{\mu-\sqrt{3}\sigma}^{\mu+\sqrt{3}\sigma}$$

$$=\frac{1}{8\sqrt{3}\sigma}\left[\left(\mu+\sqrt{3}\sigma\right)^4-\left(\mu-\sqrt{3}\sigma\right)^4\right]=\frac{1}{8\sqrt{3}\sigma}\left(8\sqrt{3}\sigma\mu^3+24\sqrt{3}\sigma^3\mu\right)$$

$$\Rightarrow m_3=\mu^3+3\sigma^2\mu$$

مثال 18-3

بعنوان دومین مثال ، یک متغیر تصادفی که دارای تابع پخش مثلثی شکل و قرینه حول محور $x=k$ را در نظر می گیریم. فرض می کنیم احتمال های $P(x<k-q)$ و $P(x>k+q)$ صفر باشند. شکل 16-3 تابع PDF را نشان می دهد.



$$\Phi(x)=\begin{cases} 0 & \text{برای } x \le k-q \\ \dfrac{h}{q}(x-k+q) & \text{برای } k-q \le x \le k \\ \dfrac{h}{q}(-x+k+q) & \text{برای } k \le x \le k+q \\ 0 & \text{برای } x \ge k+q \end{cases}$$

یا بطور اختصار

$$\Phi(x)=\begin{cases} \dfrac{h}{q}(q-|x-k|) & \text{برای } |x-k| \le q \\ 0 & \text{برای } |x-k| \ge q \end{cases}$$

ملاحظه می شود که تابع مثلثی فوق دارای همان پارامترهای (k,q,h) تابع مستطیل شکل مثال 17-3 می باشد. دراینجا نیز پارامتر h را با اعمال شرط $\int_{-\infty}^{+\infty}\Phi(x)dx$ حذف می کنیم.

این انتگرال ها همان مساحت مثلث در شکل فوق می باشند. بنابراین داریم

$$\frac{1}{2}(2q)(h)=qh=1 \Rightarrow h=\frac{1}{q}$$

تابع $\Phi(x)$ بصورت زیر در می آید

$$\Phi(x)=\begin{cases} \dfrac{1}{q}-\dfrac{|x-k|}{q^2} & \text{برای } |x-k| \le q \\ 0 & \text{برای } |x-k| \ge q \end{cases}$$

روش محاسبه برای تعیین میانگین و واریانس تابع PDF مثلثی شکل فوق همان است که در مثال 17-3 برای تابع PDF مستطیل شکل بکار بردیم. ما در اینجا نتایج را بیان می کنیم و اثبات صحت آنها را به عهده دانشجویان است.

$$\mu = k$$

$$\sigma^2 = \frac{1}{6}q^2$$

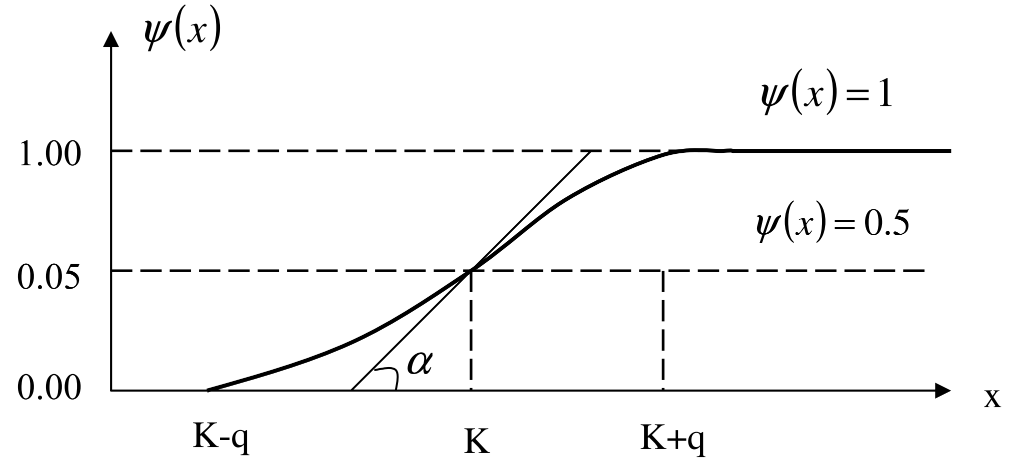
از آنجائیکه $q = \sqrt{6}\sigma$ و $k = \mu$ می باشد. می توانیم تابع PDF مثلثی شکل را که با T نشان می دهیم برحسب میانگین و واریانس تابع نشان می دهیم.

$$T(\mu,\sigma;x) = \Phi(x) = \begin{cases} \dfrac{1}{\sqrt{6}\sigma} - \dfrac{|x-\mu|}{6\sigma^2} & \text{برای } |x-\mu| \le \sqrt{6}\sigma \\ 0 & \text{برای } |x-\mu| \ge \sqrt{6}\sigma \end{cases}$$

تابع CDF مربوطه بصورت زیر خواهد بود :

$$\psi(x) = \begin{cases} 0 & \text{برای } x \le \mu - q \\ \displaystyle\int_{\mu-q}^{x} \left(\frac{1}{q} - \frac{|x-\mu|}{q^2}\right) dx & \text{برای } |x-\mu| \le q \\ 1 & \text{برای } x \ge \mu + q \end{cases}$$

تابع فوق در شکل زیر نشان داده شده است



شکل 17-3

انتگرال معادله فوق می تواند بصورت زیر هم نوشته شود

$$\int_{k-q}^{x}\left(\frac{1}{q}-\frac{|x-\mu|}{q^2}\right)dx = \begin{cases} \displaystyle\int_{\mu-q}^{x}\frac{q+x-\mu}{q^2}\,dx & \text{برای } (x\le\mu) \\[2ex] \displaystyle\int_{\mu-q}^{x}\frac{q+x-\mu}{q^2}\,dx+\int_{\mu}^{x}\frac{q-x+\mu}{q^2}\,dx & \text{برای } (x\ge\mu) \end{cases}$$

انتگرال های فوق بصورت زیر محاسبه می شوند

$$\frac{1}{q^2}\int_{\mu-q}^{x}(q+x-\mu)dx=\frac{1}{q^2}\left\{\frac{1}{2}\left[x^2\right]_{\mu-q}^{x}+(q-\mu)[x]_{\mu-q}^{x}\right\}$$

$$=\frac{1}{q^2}\left\{\frac{1}{2}\left(x^2-\mu^2+2\mu q-q^2\right)+(q-\mu)(x-\mu+q)\right\}$$

$$=\frac{1}{2q^2}\left\{x^2-2\mu x+\mu^2+2q(x-\mu)+q^2\right\}$$

$$=\frac{(x-\mu)^2}{2q^2}+\frac{(x-\mu)}{q}+\frac{1}{2}$$

همینطور

$$\frac{1}{q^2}\int_{\mu}^{x}(q-x+\mu)dx=-\frac{(x-\mu)^2}{2q^2}+\frac{(x-\mu)}{q}$$

$$\int_{\mu-q}^{x}\frac{(q+x-\mu)}{q^2}dx=\frac{1}{2}$$

در خاتمه تابع CDF را که با T_c نشان می دهیم برحسب σ,μ می نویسیم :

$$T_c(\mu,\sigma;x)=\psi(x)=\begin{cases} 0 & \text{برای } x\le\mu-\sqrt{6}\sigma \\[1ex] \dfrac{(x-\mu)^2}{12\sigma^2}+\dfrac{(x-\mu)}{\sqrt{6}\sigma}+\dfrac{1}{2} & \text{برای } \mu-\sqrt{6}\sigma\le x\le\mu \\[2ex] -\dfrac{(x-\mu)^2}{12\sigma^2}+\dfrac{(x-\mu)}{\sqrt{6}\sigma}+\dfrac{1}{2} & \text{برای } \mu\le x\le\mu+\sqrt{6}\sigma \\[1ex] 1 & \text{برای } x\ge\mu+\sqrt{6}\sigma \end{cases}$$

با روش های مشابه با مثال 17-3 می توان احتمال های $P(\mu - q \le x \le \mu + q)$ و $P(\mu - 2q \le x \le \mu + 2q)$ و $P(\mu - 3q \le x \le \mu + 3q)$ و ممان سوم m_3 تابع مثلثی شکل را پیدا کرد. در زیر نتایج بدست آمده درج شده است.

$$P(\mu - q \le x \le \mu + q) \cong 0.66$$
$$P(\mu - 2q \le x \le \mu + 2q) \cong 0.97$$
$$P(\mu - 3q \le x \le \mu + 3q) = 1$$
$$m_3 = \mu^3 + 3\sigma^2\mu$$

اثبات صحت روابط فوق بعهده دانشجو می باشد.

3-3 چند متغیر تصادفی

3-3-1 چند متغیری و توابع PDF و CDF آن

با ایده ای که در بخش 1-2-3 در مورد تابع تصادفی و متغیر تصادفی داشتیم مبحث تابع تصادفی چند ارزشی

$$X \in \{U \to R^s\}$$

در فضای s بعدی را معرفی می کنیم .

ملاحظه می شود که X یک تابع برداری یعنی X می تواند به صورت زیر نوشته شود:

$$X(u) = \left(x^1(u), x^2(u), \cdots, x^s(u)\right) \in R \qquad , \qquad u \in U$$

اجزاء $x^j(u) \in R, j=1,2,\ldots,s$ را مؤلفه های تابع $X(u)$ می نامند.

ملاحظه می شود که هر یک از مؤلفه های ، x^j ، تابع $X(u)$ را می توان بـه تنهـائی یـک متغیـر تصادفی (بخش 1-2-3) در نظر گرفت . یک مقدار بخـصوص از متغیـر x^j را ممکـن اسـت بـا x_i^j * نشان داد و همینطور یک مقدار بخصوص از تابع X را با X_i نشان داد .

$$X_i = (x_i^1, x_i^2, \ldots, x_i^S)$$

ملاحظه می شود که یک مقدار بخصوصی از تابع X عبارتست از یـک ردیـف از اعـداد حقیقـی (مثل یک مجموعه) و یا یک بردار عددی .

یک زوج ($X(u)$, $\Phi(X)$) را جائیکه

$$\Phi(X) = \Phi(x^1, x^2, \ldots, x^S) \in \{R^S \to R\} \qquad (3\text{-}27)$$

یک تابع غیر منفی و قابل انتگرال گیری روی R^S می باشد یک چند متغیری تصادفی ویـا بطـور ساده چند متغیری می نامند .

* اندیس های بالا و پایین i، و j در اینجا وسیله مفیدی هستند در تشخیص بین یک مولفه $j=1,2,\ldots,s$ وx^j وx^j_i $i=1,2,\ldots,n_j$ از چندمتغیری X و یک عنصر بخصوص یک متغیری x^j .

می توانیم از احتمال فاصله $X \in [X_0, X_1] \subset R^S$ که بصورت زیر تعریف می شود صحبت کنیم

$$P(X_0 \le X \le X_1) = \int_{X_0}^{X_1} \Phi(X)dX \in [0,1]$$

در اینجا علامت $\int$ نماینده یک انتگرال چندگانه یعنی انتگرال در فضای s بعدی می باشد و dX یک دیفرانسیل s بعدی است $dX = (dx^1, dx^2, \ldots, dx^S)$ و $X_0 = (x_0^1, x_0^2, \ldots, x_0^S)$ و $X_1 = (x_1^1, x_1^2, \ldots, x_1^S)$ حد پایینی و بالای انتگرال هستند. بطوریکه

$$X_1^j \ge X_0^j \quad \text{و} \quad j = 1,2,\ldots,S$$

برای اینکه بتوان تابع Φ را یک تابع احتمال نامید ، باید شرط زیر بر قرار باشد

$$\int_{R^S} \Phi(X)dX = 1 \tag{3-29}$$

تعریف تابع CDF در اینجا نیز کاملاً مشابه تعریف آن در حالت یک متغیری است یعنی :

$$\varphi(X) = \int_{-\infty}^{X} \Phi(Y)dY \in \{R^S \to R\} \tag{3-30}$$

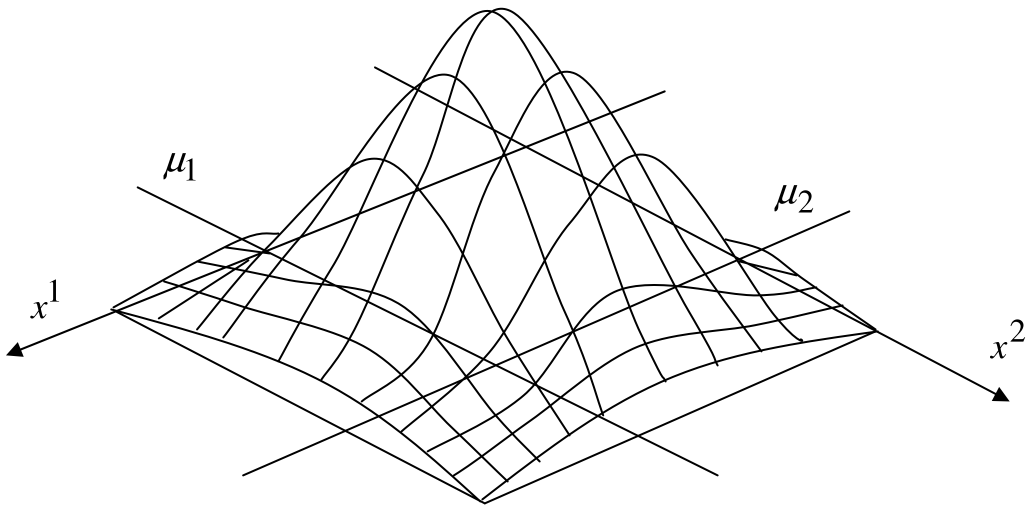
جائیکه Y متغیر زیر انتگرال در فضای s بعدی است

مثال 3-19

تابع یک متغیری در شکل 3-12 را در نظر بگیرید . این تابع به شکل زنگ بنام تابع پخش نرمال و یا گوس ، که بعداً بحث خواهد شد ، مشهور است و آنرا معمولاً با N بر حسب میانگین و انحراف معیار آن به صورت زیر نشان می دهند.

$$\varphi(X) = N(\mu, \sigma; x)$$

تابع PDF یک دو متغیری نرمال $\Phi(x)=\Phi(x^1,x^2)$ بشکل زیر ظاهر می شود



(شکل 3-18)

در فضای دو بعدی تابع $\Phi(x)$ را تابع پخش دو متغیری می نامند. تابع دو متغیری نرمال فوق به صورت زیر نشان داده می شود.

$$\varphi(X) = N(\mu_1, \mu_2, \sigma_1, \sigma_2; x)$$

3-3-2 وابستگی و استقلال آماری

تابع پخش PDF یک چند متغیری ممکن است دارای شکل زیر باشد

$$\Phi(x) = \Phi_1(x^1)\Phi_2(x^2)...\Phi_s(x^s) = \prod_{j=1}^{s} \Phi_j(x^j)$$

در این صورت انتگرال معادله 28-3 به شکل زیرنوشته خواهد شد.

$$\int_{X_\circ}^{X_1} \Phi(x)dx = \int_{X_\circ}^{X_1} \prod_{j=1}^{s} \Phi_j(x^j)dx^j = \prod_{j=1}^{s} \int_{X_\circ^j}^{X_1^j} \Phi_j(x^j)dx^j \qquad (3\text{-}31)$$

با در نظر گرفتن اینکه هر مؤلفه x^j از چند متغیری X خود یک متغیر تصادفی ا تابع پخش Φ_j می باشد معادله 31-3 را می توان به صورت زیر نوشت

$$\prod_{j=1}^{s} \int_{X_\circ^j}^{X_1^j} \Phi_j(x^j)dx^j = \prod_{j=1}^{s} P\left(X_\circ^j \le X^j \le X_1^j\right)$$

از مقایسه نتیجه فوق با معادله 28-3 می توان نوشت

$$P(X_\circ \le X \le X_1) = \prod_{j=1}^{s} P\left(X_\circ^j \le X^j \le X_1^j\right) \qquad (3\text{-}32)$$

رابطه فوق چنین تعبیر می شود که احتمال مرکب تمام مؤلفه ها در فواصل $[x^j_0, x^j_1]$ برابر است با حاصلضرب تک تک احتمال ها. مسلماً چنین احتمال مرکبی در مورد چند حادثه مستقل صادق است. بنابراین مؤلفه های x^j از چند متغیری X که دارای تابع پخش $\Phi(x)$ تعریف شده در فوق هستند از نظر آمار مستقل از همدیگر می باشند.

اگر تابع احتمال یک چند متغیری بتواند به صورت حاصلضرب توابع احتمال مؤلفه ها نوشته شود می گویند که آن مؤلفه ها از نظر آمار وابسته هستند. در این صورت احتمال $P(X_0 \le X \le X_1)$

مساوی حاصلضرب احتمال مولفه ها نخواهد بود. در مورد مولفه های مستقل از همدیگر می توان نوشت

$$\int_R \Phi_j\left(x^j\right)dx^j = 1 \qquad \text{و} \qquad j = 1,2,\ldots,S$$

3-3-3 میانگین و واریانس یک چند متغیری

بردار

$$\tilde{\mu} = (\mu_1, \mu_2, \ldots, \mu_s) = \tilde{E}^*(X) \in R^S \tag{3-33}$$

را که در آن

$$\mu_j = \int_{R^S} x^j \Phi(x)dx = E^*\left(X^j\right) \in R \qquad \text{و} \qquad j = 1,2,\ldots,S \tag{3-34}$$

باشد میانگین چند متغیری X می نامند.

اپراتور $\tilde{E}^*$ (انتگرال در فضای S بعدی) روی

$$X\Phi(X) = \left(x^1, x^2, \ldots, x^s\right)\Phi\left(x^1, x^2, \ldots, x^s\right)$$

اعمال می شود.

واریانس یک چند متغیری X طبق رابطه زیر تعریف می شود

$$\tilde{\sigma}^2 = \left(\sigma_1{}^2, \sigma_2{}^2, \ldots, \sigma_s{}^2\right) = \tilde{E}^*\left((X - \tilde{\mu})^2\right) \in R^S \tag{3-35}$$

که در آن

$$\sigma_j{}^2 = \int_{R^S} \left(x^j - \mu_j\right)^2 \Phi(x)dx$$

$$\tilde{E}^*\left(x^j - \mu_j\right)^2 \in R \qquad \text{و} \qquad j = 1,2,\ldots,S \tag{3-36}$$

واریانس بر حسب اپراتور $\tilde{E}^*$ به صورت زیر بسط داده می شود

$$\tilde{E}^*\left((X - \tilde{\mu})^2\right) = \tilde{E}^*\left(\left(X - \tilde{E}^*(X)\right)^2\right) = \tilde{E}^*\left(X^2\right) - \tilde{\mu}^2 \tag{3-37}$$

و

(3-38)

$$E^*\left(\left(X^j - \mu_j\right)^2\right) = E^*\left(\left(X^j - E^*\left(X^j\right)\right)^2\right) = E^*\left(\left(X^j\right)^2\right) - \mu_j{}^2$$

واریانس یک چند متغیری نمی تواند بطور کامل خصوصیات آماری آنرا در فضای چند بعدی بیان کند بر عکس حالت یک متغیری که در آنجا واریانس یک متغیر تصادفی بیانگر خصوصیات کامل آماری آن در فضای یک بعدی می باشد. بنابراین برای تکمیل مشخصات آماری یک چند متغیری می پردازیم به معرفی باصطلاح ماتریس واریانس کوواریانس (بخش 3-3-4) .

قبل از صحبت درباره ماتریس فوق میانگین و واریانس یک چند متغیری را که مؤلفه های آن $x^j, j=1,2,...,S$ مستقل از همدیگر می باشند بررسی می کنیم. میانگین مؤلفه های مستقل عبارتند از

$$\mu_j = \int_{R^S} x^j \prod_{\ell=1}^{S} \Phi_\ell\left(x^\ell\right) dx^\ell = \int_{R^S} x^j \Phi_j\left(x^j\right) dx^j \prod_{\substack{\ell=1 \\ \ell \neq j}}^{S} \Phi_\ell\left(x^\ell\right) dx^\ell$$

$$= \int_R x^j \Phi_j\left(x^j\right) dx^j . \prod_{\substack{\ell=1 \\ \ell \neq j}}^{S} \int_R \Phi_\ell\left(x^\ell\right) dx^\ell \qquad (3\text{-}39)$$

تمام انتگرال های بعد از علامت $\prod$ در فوق طبق بخش 3-3-2 برابر واحد هستند. بنابراین داریم

$$\mu_j = \int_R x^j \Phi_j\left(x^j\right) dx^j \qquad (3\text{-}40)$$

همینطور برای واریانس خواهیم داشت

$$\sigma_j^{\,2} = \int_R \left(x^j - \mu_j\right)^2 \Phi_j\left(x^j\right) dx^j \qquad \text{و} \qquad j = 1,2,...,S \qquad (3\text{-}41)$$

بنابراین برای یک چند متغیری که دارای مؤلفه های مستقل از همدیگر می باشد ، میانگین و واریانس هر مؤلفه می تواند بطور جداگانه محاسبه شود.

3-3-4 کوواریانس و ماتریس واریانس کوواریانس

قبل از اینکه ماتریس واریانس کوواریانس مورد بحث قرار گیرد احتیاج است که یکی دیگر از کمیتهای آماری در اینجا گفته شودو آن کوواریانس می باشد. برای دو مؤلفه x^k, x^j از یک چند متغیری کوواریانس مطابق زیر تعریف می شود

$$\operatorname{cov}\left(x^k, x^j\right) = \sigma_{jk} = \int_{R^S} \left(x^j - \mu_j\right)\left(x^k - \mu_k\right)\Phi(x)dx$$

$$= E^*\left(x^j - \mu_j\right)\left(x^k - \mu_k\right)\Phi(x)dx = \sigma_{kj} \in R; k, j = 1,2,...,S \qquad (3\text{-}42)$$

سه چیز در معادله فوق مشاهده می شود. اول اینکه برای $j=k$ رابطه فوق تبدیل به واریانس می شود. یعنی

برای $j=k$ $\qquad \sigma_{jk}=\sigma_{kj}={\sigma_j}^2$

دوم اینکه اگر مؤلفه های یک چند متغیری مستقل از همدیگر باشند کوواریانس دوبدو از آنها $j\neq k$ همه مساوی صفر هستند. برای نشان دادن آن می دانیم

$$\sigma_{jk}=\int_{R^S}\left(x^j-\mu_j\right)\left(x^k-\mu_k\right)\prod_{\ell=1}^{s}\Phi_\ell\left(x^\ell\right)dx^\ell$$
$$=\int_R\left(x^j-\mu_j\right)\Phi_j\left(x^j\right)dx^j\int_R\left(x^k-\mu_k\right)\Phi_k\left(x^k\right)dx^k$$
$$=\left[\int_R x^j\Phi_j\left(x^j\right)dx^j-\mu_j\right]\left[\int_R x^k\Phi_k\left(x^k\right)dx^k-\mu_k\right]$$
$$\left[\mu_j-\mu_j\right]\left[\mu_k-\mu_k\right]=0$$

سوم اینکه با توجه به کوواریانس یک زوج متغیر مستقل که مساوی صفر می باشند یعنی

$$\sigma_{jk}=E^*\left(\left(X^j-\mu_j\right)\left(X^k-\mu_k\right)\right)=0 \qquad (3\text{-}43)$$

می توان نوشت

$$\begin{aligned}\sigma_{jk}&=E^*\left(\left(X^j-\mu_j\right)\left(X^k-\mu_k\right)\right)=E^*\left(X^jX^k-\mu_kX^j-\mu_jX^k+\mu_j\mu_k\right)\\&=E^*\left(X^jX^k\right)-\mu_kE^*\left(X^j\right)-\mu_jE^*\left(X^k\right)+\mu_j\mu_k\\&=E^*\left(X^jX^k\right)-2\mu_j\mu_k+\mu_j\mu_k=E^*\left(X^jX^k\right)-\mu_k\mu_j=0\end{aligned}$$

بنابراین برای مؤلفه های مستقل x^k,x^j خواهیم داشت

$$E^*\left(X^jX^k\right)=E^*\left(X^j\right)E^*\left(X^k\right) \qquad (3\text{-}44)$$

با تعمیم رابطه فوق برای r متغیر مستقل خواهیم داشت

$$E^*\left(\prod_{j=1}^{r}X^j\right)=\prod_{j=1}^{r}E^*\left(X^j\right) \qquad (3\text{-}45)$$

رابطه 3-45 لیست خصوصیات اپراتور E^* را در بخش 3-2-3 تکمیل می کند.

بطوریکه در بخش 3-2-3 گفته شد بردار واریانس $\tilde{\sigma}^2$ به تنهایی کافی برای بیان کامل خصوصیات آماری یک چند متغیری ، تا حدود ممانهای دوم ، نیست. برای بدست آوردن همان مقدار اطلاعات آماری که در حالت یک متغیری از واریانس حاصل می شد باید کوواریانس ها

را نیز در نظر گرفت. واریانس ها و کوواریانس ها با هم در یک ماتریس بنام ماتریس واریانس کوواریانس و یا فقط ماتریس کوواریانس گردآوری می شوند.

ماتریس واریانس کوواریانس یک چند متغیری X معمولا با $\sum_X{}^*$ و بشکل زیر نشان داده می شود.

$$\sum_X{}^* = \begin{bmatrix} \sigma_1{}^2 & \sigma_{12} & \sigma_{13} & \ldots & \sigma_{1S} \\ \sigma_{21} & \sigma_2{}^2 & \sigma_{23} & \ldots & \sigma_{2S} \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ \sigma_{S1} & \sigma_{S2} & \sigma_{S3} & \ldots & \sigma_S{}^2 \end{bmatrix} \tag{3-46}$$

بسادگی دیده می شود که ماتریس واریانس کوواریانس می تواند برحسب اپراتور E^* بشکل زیر نوشته شود

$$\sum_X{}^* = \tilde{E}^*\left[\left(X - \tilde{E}^*(X)\right)\left(X - \tilde{E}^*(X)\right)^T\right] \tag{3-47}$$

توان T در فوق علامت ترانسپوز ماتریس می باشد. اثبات رابطه فوق بعهده دانشجویان می باشد.

ملاحظه می شود که ماتریس واریانس کوواریانس همیشه یک ماتریس قرینه است. عناصر روی قطر اصلی عبارتند از واریانس مؤلفه و عناصر باقیمانده کوواریانس زوج های مختلف از مؤلفه های چند متغیری می باشند. شرط لازم و کافی برای اینکه ماتریس واریانس کوواریانس قطری باشد یعنی عناصر خارج از قطر اصلی صفر باشند اینست که مؤلفه های چند متغیری مستقل از همدیگر باشند. ماتریس واریانس کوواریانس یکی از اساسی ترین کمیتها در محاسبات سرشکنی است.

3-3-5 چند نمونه ای تصادفی و توابع PDF و CDF آن

نظیر حالت یک متغیری ، کمیت η مربوط به نمونه تصادفی ξ (بخش 3-3-1) را مطابق زیر تعریف می کنیم

$$\eta = \begin{bmatrix} \xi^1 \\ \xi^2 \\ . \\ . \\ . \\ \xi^S \end{bmatrix} = \begin{bmatrix} \left(\xi_1^1, \xi_2^1, ..., \xi_{n_1}^1\right) \in R^{n1} \\ \left(\xi_1^2, \xi_2^2, ..., \xi_{n_2}^2\right) \in R^{n2} \\ . \\ . \\ . \\ \left(\xi_1^S, \xi_2^S, ..., \xi_{n_S}^S\right) \in R^{nS} \end{bmatrix} \quad (3\text{-}48)$$

کمیت η ار تعمیم بلافاصله نمونه تصادفی ξ حاصل شده و چند نمونه ای تصادفی حاصل می شود. از تعریف فوق چنین برمی آید که η دارای S مؤلفه ، ξ^j ، است که هر کدام از آنها به تنهائی یک نمونه تصادفی می باشند. تعداد عناصر هر مؤلفه ، n_j ، ممکن است برای تمام مؤلفه ها یکسان و یا متفاوت باشد.

همچنین می توان مجموعه تعاریف و تابع PDF و CDF یک چند نمونه ای را درست نظیر یک نمونه تصادفی تعریف کرد. همینطور نمودارها و پلی گن های تجمعی و غیر تجمعی می توانند برای یک دو نمونه ای مورد استفاده قرار گیرند.

3-3-6 میانگین و ماتریس واریانس کوواریانس یک چند نمونه ای

میانگین یک چند نمونه ای بشکل زیر تعریف می شود

$$\tilde{\mu} = \left(M_1, M_2, ..., M_S\right) = \tilde{E}(\eta) \in R^S \quad (3\text{-}49)$$

بطوریکه با استفاده از فرمول 3-3 داریم

$$\mu_j = \frac{1}{n_j}\sum_{i=1}^{n_j}\xi_i^j = E\left(\xi^j\right) \in R \qquad \text{و} \qquad j = 1,2,...,S \quad (3\text{-}50)$$

در اینجا اپراتور $\tilde{E}$ عبارتست از یک بردار از اپراتورهای E . این را می توان از مقایسه دو فرمول 3-50و3-49 در یافت.

واریانس یک چند نمونه ای به شکل زیر تعریف می شود.

$$\tilde{S}^2 = \left(S_1{}^2, S_2{}^2, ..., S_s{}^2\right) = \tilde{E}\left(\eta - \tilde{M}\right)^2 \in R^s \quad (3\text{-}51)$$

به طوری که از معادله 3-6 داریم

$$S_j^2 = \frac{1}{n_j}\sum_{i=1}^{n}(\xi_i^j - M_j)^2 = E(\xi^j - M_j)^2 \in R,\ j = 1,2,...,S \quad (3\text{-}52)$$

بردار انحراف معیار S طبق رابطه زیر تعریف می شود

$$\tilde{S} = (S_1, S_2, ..., S_s) \qquad (3\text{-}53)$$

مثال 3-20

مطلوب است تعیین میانگین $\tilde{M}$ و واریانس $\tilde{S}^2$ و انحراف معیار $\tilde{S}$ نمونه $\eta=(\zeta^1,\zeta^2,\zeta^3)$ به طوری که

$$\zeta^1=(2,3,4,7,4)$$
$$\zeta^2=(6,4,0,3,2)$$
$$\zeta^3=(5,2,5,5,8)$$

در اینجا داریم $n_1=n_2=n_3=5$. بردار $\tilde{M}$ از معادله 3-49 محاسبه می شود

$$\tilde{M}=(M_1,M_2,M_3)$$

$$M_1 = \frac{1}{n_1}\sum_{i=1}^{m_1}\xi_i^1 = \frac{1}{5}\sum_{1}^{5}\xi_i^1 = \frac{1}{5}(2+3+4+7+4) = \frac{20}{5} = 4$$

$$M_2 = \frac{1}{n_2}\sum_{i=1}^{n_2}\xi_i^2 = \frac{1}{5}(6+4+0+3+2) = \frac{15}{5} = 3$$

$$M_3 = \frac{1}{5}(5+2+5+5+8) = \frac{25}{5} = 5$$

بنابراین داریم $\tilde{M}=(4,3,5)$

برای محاسبه واریانس از فرمول 3-51 استفاده می کنیم.

$$\tilde{S} = (S_1^2, S_2^2, S_3^2)$$

$$S_1^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(\xi_i^1 - M_1)^2 = \frac{1}{5}\sum_{i=1}^{5}(\xi_i^1 - 4)^2 = \frac{1}{5}[(-2)^2 + (-1)^2 + (0)^2 + (3)^2 + (0)^2] = 2.8$$

$$S_2^2 = \frac{1}{5}[(4)^2 + (1)^2 + (-3)^2 + (0)^2 + (-1)^2] = 4.0$$

$$S_3^2 = \frac{1}{5}[(0)^2 + (-2)^2 + (0)^2 + (0)^2 + (3)^2] = 3.6$$

$$\tilde{S^2} = (2.8, 4.0, 3.6)$$

جذر هر کدام از S^2_j ها یعنی S_j ها بردار $\tilde{S}$ را تشکیل می دهند.

$$\tilde{S} = (S_1, S_2, S_3) = (1.67, 2.0, 1.9)$$

اگر مولفه های ζ^j و ζ^k از چند نمونه ای η دارای تعداد مساوی عناصر باشد $(n_i=n_k=n)$،

کوواریانس S_{jk} بین دو مولفه مزبور طبق رابطه زیر تعریف می شود

$$S_{jk} = \frac{1}{n}\sum_{i=1}^{n}[(\xi_i^j - M_j)(\xi_i^k - M_k)] \qquad (3\text{-}54)$$

رابطه فوق بر حسب اپراتور E به صورت زیر نوشته می شود

$$S_{jk} = E\left((\xi^j - M_j)(\xi^k - M_k)\right) = S_{kj} \in R \quad \text{و} \quad j,k = 1,2,...,s \qquad (3\text{-}55)$$

ملاحظه می شود که مقدار S_{jk} بستگی به ترتیب عناصر در هر دو نمونه ζ^j و ζ^k دارد. در صورتی که میانگینهای M_j وM_k و واریانس S^2_j وS^2_k اینطور نیستند.

بنابراین برای محاسبه کوواریانس، عناصر هر دو نمونه ζ^j و ζ^k باید به همان ترتیبی که بدست آمده اند منظور شوند. مثال زیر این مسئله را روشن می کند. فرض کنید که عناصر ζ^j عبارتند از مشاهدات یک زاویه قائم و عناصر ζ^k عبارتند از زمانهای مشاهده زوایای قائم. مسلما برای مطالعه رابطه (کوواریانس) بین زمان مشاهده ومقدار زاویه قائم باید هر زاویه کنار زمان مربوطه در فرمول 54-3 قرار می گیرد.

مثال 21-3

مطلوب است تعیین کوواریانس بین زوج های مختلف از مولفه های چند نمونه ای η داده شده در مثال 20-3، برای این کار از فرمول 54-3 استفاده می کنیم

$$S_{12} = S_{21} = \frac{1}{5}\sum_{i=1}^{5}[(\xi_i^1 - 4)(\xi_i^1 - 3)]$$

$$= \frac{1}{5}[(-2)(3) + (-1)(1) + (0)(-3) + (3)(0) + (0)(-1)] = -1.4$$

$$S_{13} = S_{31} = \frac{1}{5}[(-3)(0) + (-1)(-3) + (0)(0) + (3)(0) + (0)(3)] = 0.6$$

$$S_{23} = S_{32} = \frac{1}{5}[(3)(0) + (1)(-3) + (-3)(0) + (0)(0) + (-1)(3)] = -1.2$$

در خاتمه، ماتریس واریانس-کوواریانس Σ_η چند نمونه ای را به صورت زیر گرد آوری می شود.

$$\sum \eta = \begin{bmatrix} S_1{}^2 & S_{12} & S_{13} & \dots & S_{1S} \\ S_{21} & S_2{}^2 & S_{23} & \dots & S_{2S} \\ S_{31} & S_{32} & S_3{}^2 & \dots & S_{3S} \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ S_{S1} & S_{S2} & S_{S3} & \dots & S_S{}^2 \end{bmatrix} \qquad (3\text{-}56)$$

تا اینجا میانگین و ماتریس واریانس- کوواریانس یک چند نمونه ای را شناختیم. در بخش 3-3-3 گفتیم که بسط از یک بعدی به S بعدی عبارت از فهم مستقیم واریانس یک بعدی می باشد. برای بیان خصوصیات آماری یک چند متغیری سطح ممانهای دوم مجبور بودیم ماتریس واریانس-کوواریانس را معرفی کنیم. اگر رابطه بین نمونه و یک متغیر را در نظر بگیریم، این رابطه بین چند نمونه ای و چند متغیری برقرار نیست. فرمول های میانگین و

واریانس یک نمونه معادل فرمول های مربوطه برای یک متغیری است در حالیکه فرمول های مزبور برای چند نمونه ای و چند متغیری معادل همدیگر نیستند. فرمول هایی نظیر فرمول های 3-34 و3-35 و3-42 را می توان برای یک چند نمونه ای طرح نمود. فرمول هایی که بیشتر در عمل برای چند نمونه ای مورد استفاده قرار می گیرند عبارتند از 3-49و3-51و3-54. این فرمول ها در واقع نظیر فرمول های 3-40و3-41و3-43 که تنها برای چند متغیری مستقل مورد استفاده قرار می گیرند هستند.

با در نظر گرفتن موارد فوق ، مشکلات ناشی از محاسبه کوواریانس یک چند نمونه ای که برای آن باید تعداد عناصر موجود در هر دو نمونه یکسان باشد ، عمل یک ماتریس واریانس-کوواریانس برای چند نمونه ای فرض می کنیم. انتخاب چنین ماتریسی برای چند نمونه ای یکی از کارهای حساس در محاسبات سر شکنی می باشد.

مثال 3-22

مطلوب است تعیین ماتریس واریانس-کوواریانس چند نمونه ای η در مثال 3-20. واریانس های مربوط به η در مثال 3-20 محاسبه شده اند و عبارتند از

$$S_1^2 = 2.8 \quad و \quad S_2^2 = 4.0 \quad و \quad S_3^2 = 3.6$$

و همینطور کوواریانس های η در مثال 3-21 محاسبه شده اند و عبارتند از

$$S_{12} = S_{21} = -1.4 \quad و \quad S_{13} = S_{31} = 0.6 \quad و \quad S_{23} = S_{32} = -1.2$$

بنابراین ماتریس واریانس-کوواریانس عبارت خواهد بود از

$$\sum \eta = \begin{bmatrix} 2.8 & -1.4 & 0.6 \\ -1.4 & 4.0 & -1.2 \\ 0.6 & -1.2 & 3.6 \end{bmatrix}$$

3-3-7 وابستگی (Correlation)

هر چند که کوواریانس های یک چند نمونه ای نقش یکسان با کوواریانس های یک چند متغیری بازی نمی کنند، اما می توان از آنها به عنوان معیاری برای سنجش وابستگی آماری استفاده کرد. می گوییم که آنها درجه وابستگی مؤلفه ها را نشان می دهند.

درجه وابستگی به عنوان معیار سنجش وابستگی آماری ممکن است منحصر به فرد نباشد.

می توان دید که کوواریانس $S_{jk} \in R$ ممکن است هر مقداری داشته باشد. بنابراین نمی تواند یک وسیله خوب برای سنجش وابستگی باشد علاوه بر آن نمی توان مقدار آن را برای حالت ماکزیمم با وابستگی کامل پیش بینی کرد.

به این دلیل معیار دیگری، غیر از کوواریانس ، به نام ضریب وابستگی که با ρ نشان داده شده به صورت زیر تعریف می کنیم

$$\rho_{jk} = \frac{S_{jk}}{S_j S_k} \qquad \text{و} \qquad S_j, S_k \neq 0 \qquad (3\text{-}57)$$

می توان نشان داد که مقدار ρ_{jk} از 1- تا 1+ تغییر می کند.

محاسبات وابستگی که یکی از شاخه های آمار می باشد بر اساس استفاده از ضریب وابستگی است.

دو مؤلفه ζ^j و ζ^k از یک چند نمونه ای ممکن است دارای حالت های زیر باشند

1. کاملا مستقل از همدیگر اگر $\rho_{jk}=0$ باشد
2. وابسته اگر $|\rho_{jk}|<1$ باشد
3. کاملا وابسته بطور مثبت اگر $\rho_{jk}=1$ باشد
4. کاملا وابسته بطور منفی اگر $\rho_{jk}=-1$ باشد

ملاحظه کنید که برای چند متغیری نیز فرمولی نظیر 57-3 برای ضریب وایستگی آن می توان نوشت.

مثال 23-4

مطلوب است تعیین درجه وابستگی بین مؤلفه های مختلف چند نمونه ای η مورد بحث در مثال های 20-3 تا 22-3. ضریب وابستگی از فرمول 57-3 به صورت زیر محاسبه می شود :

$$\rho_{12} = \frac{S_{12}}{S_1 S_2} = \frac{-1.4}{1.67 \times 2} = -0.42$$

$$\rho_{13} = \frac{S_{13}}{S_1 S_3} = \frac{0.6}{1.67 \times 1.9} = 0.19 \qquad \text{و} \qquad \rho_{32} = \frac{-1.2}{2 \times 1.9} = -0.31$$

و داریم $\rho_{12}=\rho_{21}$ و $\rho_{13}=\rho_{31}$ و $\rho_{23}=\rho_{32}$

از آنجایی که مقادیر فوق در نامساوی زیر صدق می کنند یعنی

$$\left|\rho_{jk}\right| < 1 \qquad \text{و} \qquad j,k = 1,2,3 \qquad j \neq k$$

بنابراین سه مولفه ξ_1و ξ_2 و ξ_3 ازچند نمونه ای η همگی وابسته می باشند.

مثال 24-3

مطلوب است تعیین درجه وابستگی بین مؤلفه های ξ^1 و ξ^2 مؤلفه های ξ^1 و ξ^3 از چند نمونه ای $\eta = \left(\xi^1, \xi^2, \xi^3\right)$ با مؤلفه های

$$\xi^1 = (2,1,3,5,4)$$
$$\xi^2 = (4,2,6,10,8)$$
$$\xi^3 = (-4,-2,-6,-10,-8)$$

میانگین ها و واریانس ها و کوواریانس های سه مؤلفه فوق طبق آنچه که در مثال های 20-3 و 21-3 گذشت محاسبه شده و ازای مقادیر زیر می باشند

$M_1 = 3$ و $M_2 = 6$ و $M_3 = -6$

$S_1{}^2 = 2$ و $S_2{}^2 = 8$ و $S_3{}^2 = 8$

$S_1 = \sqrt{2}$ و $S_2 = 2\sqrt{2}$ و $S_3 = 2\sqrt{2}$

$S_{12} = 4$ و $S_{13} = -4$

در نتیجه خواهیم داشت

$$\rho_{12} = \frac{S_{12}}{S_1 S_2} = \frac{4}{\sqrt{2} \times 2\sqrt{2}} = 1 \quad \text{و} \quad \rho_{13} = \frac{S_{13}}{S_1 S_3} = \frac{-4}{\sqrt{2} \times 2\sqrt{2}} = -1$$

نتایج فوق نشان می دهد که مؤلفه های ξ^1 و ξ^2 کاملاً به طور مثبت وابسته بوده و مؤلفه های ξ^1 و ξ^3 کاملاً به طور منفی وابسته می باشند

برای سهولت در کنترل محاسبات ، اغلب ، میانگین ها ، واریانس ها و کوواریانس ها و ضرایب وابستگی مؤلفه های یک چند نمونه ای در یک جدول مرتب شده و محاسبه می شوند. جدول زیر محاسبات مربوط به دو مولفه ε^1 و ε^2 در چند نمونه ای مثال 20-3 را نشان می دهد.

	ξ^1			ξ^2			$(\xi_i^1 - M_1) \times$
	ξ_i^1	$(\xi_i^1 - M_1)$	$(\xi_i^1 - M_1)^2$	ξ_i^2	$(\xi_i^2 - M_2)$	$(\xi_i^2 - M_2)^2$	$(\xi_i^2 - M_2)$
	2	-2	4	6	3	9	-6
	3	-1	1	4	1	1	-1
	4	0	0	0	-3	9	0
	7	3	9	3	0	0	0
	4	0	0	2	-1	1	0
$\sum$	20		14	15		20	-7

$$M_1 = \frac{1}{5}(20) = 4 \qquad M_2 = \frac{1}{5}(15) = 3$$
$$S_1{}^2 = \frac{1}{5}(14) = 2.8 \qquad S_2{}^2 = \frac{1}{5}(20) = 4$$
$$S_1 = \sqrt{2.8} = 1.67 \qquad S_2 = \sqrt{4} = 2$$
$$S_{12} = \frac{1}{5}(-7) = -1.4$$

$$\rho_{12} = \frac{-1.4}{1.67 \times 2} = -0.42$$

3-4 تمرین شماره 3

1- جدول زیر یک نمونه تصادفی از معدل دانش آموزان دبیرستانی می باشد.این نمونه وزن بیست نفر را به پوند نشان می دهد.

الف-میانگین، انحراف معیار(Standard deviation) ، میانه (Medium) وعرض (Range) نمونه را با استفاده از خود نمونه و بعد به وسیله مجموعه تعاریف آن بدست آورید.

ب - احتمالات تجربی عناصر نمونه را پیدا کرده، تابعPDF وCDF آن را رسم کنید.

پ - احتمال اینکه دانش آموزی کمتر و مساوی 150 پوند داشته باشد چقدر است.

ت – احتمال اینکه وزن دانش آموزی در فاصله [158,173] باشد چقدر است.

138	150	141	158	150
146	164	138	164	164
150	146	158	173	150
158	140	146	150	164

2-جدول زیر یک نمونه تصادفی از ارتفاع های نوعی درخت می باشد. در این نمونه ارتفاع 125 درخت به سانتیمتر اندازه گیری ومرتب شده است.

61 70 102 107 113 114 114 119 120 126 126 129 129 132 137
139 149 142 143 146 146 147 147 148 149 149 150 150 153
139
156 157 158 159 159 159 159 162 162 164 166 166 166 167
153
169 169 169 170 171 171 172 172 172 173 173 175 175 176
169
176 177 177 178 179 180 180 181 181 181 182 182 184 184
176
186 187 188 188 190 192 192 193 194 194 194 195 195 195
185
197 198 198 200 201 201 201 202 202 203 204 206 206 209
196
209 214 216 219 219 219 221 222 223 227 233 234 236 237
209
246 247 254 262 270

الف- نمونه فوق را به اختیار کلاسه بندی کرده، نمودار پلی گن غیر تجمعی و تجمعی آنرا رسم کنید.
ب - میانگین، انحراف معیار، میانه و عرض را محاسبه کرده مقادیر بدست آمده را در نمودار ها و پلی گن های رسم شده وارد کنید.
پ – احتمالی را که ارتفاع درخت در فاصله [121,174] سانتیمتر باشد با استفاده از نمودارها و پلی گن های رسم شده (تجمعی و غیر تجمعی) و با استفاده از نمونه واقعی بدست آورید سپس مقادیر بدست آمده را با هم مقایسه کنید.

3-نتایج داده شده در مثال 18-3 را برای میانگین، واریانس و ممان حول صفر درجه سوم تابع PDF مثلثی شکل با استفاده از فرمول های مربوطه محاسبه کنید.

4-نتایج بدست آمده در مثال 17-3 و 18-3 برای احتمال های $P(\mu-\sigma \le x \le \mu+\sigma)$ و $P(\mu-2\sigma \le x \le \mu+2\sigma)$ و $P(\mu-3\sigma \le x \le \mu+3\sigma)$ را از طریق منحنی های مستطیل و مثلثی شکل تجمعی CDF (غیر از منحنی های PDF) کنترل کنید.

5- متغیر تصادفی x را با تابع PDF زیر در نظر می گیریم .

$$\Phi(x) = \begin{cases} h & \text{برای } (-3 \le x \le 7) \\ 0 & \text{برای } x \text{ های دیگر} \end{cases}$$

الف- مقدار h را تعیین کنید.
ب – تابع CDF آنرا رسم کنید.
پ – هر دو تابع PDF وCDF را مورد استفاده قرار داده احتمال های زیر را محاسبه کنید.
$P(x \le 1.5)$ و $P(x \ge 2.5)$ و $P(-1 \le x \le 4)$ و $P(\mu-2\sigma \le x \le \mu+2\sigma)$
ت - ممان های درجه سوم و چهارم PDF مزبور را حول صفر محاسبه کنید.

6- متغیر تصادفی x با تابع PDF زیر مورد نظر است

$$\Phi(x) = \begin{cases} kx & \text{برای } (0 \le x \le 2) \\ 0 & \text{برای } x \text{ های دیگر} \end{cases}$$

الف- میانگین، واریانس و انحراف معیار x را پیدا کنید.

ب – احتمال $P(1 \le x \le 1.5)$ را پیدا کنید.

7-متغیر تصادفی x با تابع PDF زیر را در نظر می گیریم.

$$\Phi(x) = \begin{cases} k + \frac{1}{50}x - \frac{3}{50} & \text{برای } (3 \le x \le 8) \\ k - \frac{1}{50}x + \frac{13}{50} & \text{برای } (8 \le x \le 13) \\ 0 & \text{برای x های دیگر} \end{cases}$$

الف—میانگین و انحراف معیار(Standard deviation) x را پیدا کنید.

ب – احتمال های $P(5.5 \le x \le 10.5)$ ، $P(x \le 9)$ ، $P(x \ge 7)$ و $P(\mu - \sigma \le x \le \mu + \sigma)$ را محاسبه کنید.

8-چند نمونه ای $\eta = (\xi^1, \xi^2, \xi^3)$ در زیر داده شده است.

$$\xi^1 = (4.2, 3.7, 4.1)$$
$$\xi^2 = (26.7, 26.3, 26.6)$$
$$\xi^3 = (-17.5, -17.0, -18.0)$$

الف- میانگین η را پیدا کنید.

ب – ماتریس واریانس –کوواریانس η را محاسبه کنید.

پ – تمام ضرایب وابستگی موجود بین مؤلفه ها را محاسبه کنید.

9- دو متغیری $X = (x^1, x^2)$ با تابع PDF زیر داده شده است.

$$\Phi(x) = \begin{cases} -\frac{|x^1 - q|}{s^2 t 12\sqrt{3}} + \frac{1}{st6\sqrt{2}} & \text{برای } (|x^1 - q|) \le s\sqrt{6} \text{ و } (|x^2 - r|) \le t\sqrt{3} \\ 0 & \text{در جاهای دیگر} \end{cases}$$

جاییکه q و r اعداد حقیقی هستند و s و t اعداد مثبت حقیقی هستند.

الف-میانگین X را محاسبه کنید.

ب – ماتریس واریانس-کوواریانس X را محاسبه کنید.

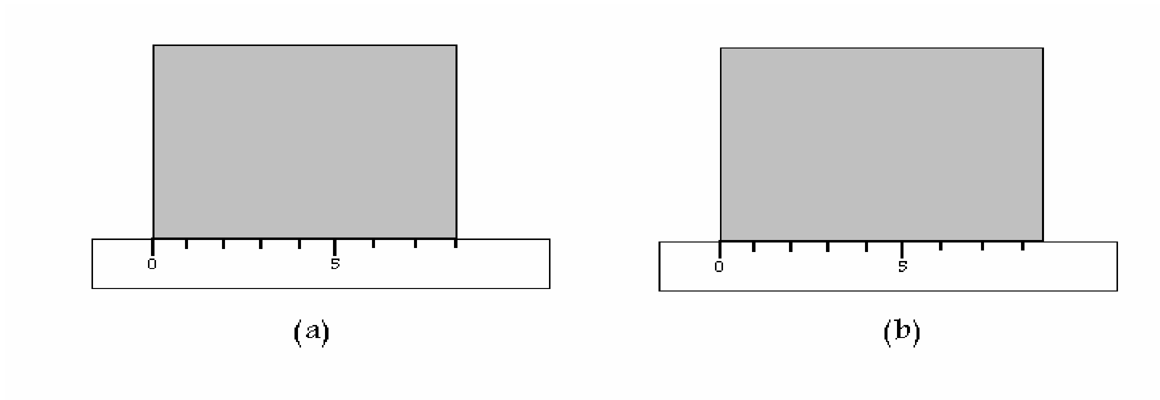
فصل 4-اساس تئوری خطاها

4-1 تعاریف اساسی

درعمل ما با مشاهدات سروکار داریم. مشاهدات چیزی نیست جز معرفی عددی کمیتهای فیزیکی نظیر طولها ، زاویه ها ، وزن ها و غیره. این مشاهدات از طریق نوعی اندازه گیری با مقایسه با استانداردهای از قبل تعریف شده بدست می آید.در اغلب حالات ما تعداد زیادی مشاهدات برای یک کمیت فیزیکی انجام می دهیم و فرض می کنیم که آنها همه معرف آن کمیت واحد هستند.

افکاری وجود دارد که بر اساس آنها هیچ کمیتی نمی تواند دوبار مورد اندازه گیری قرار گیرد.آنها عقیده دارند اگر کمیتی برای بار دوم اندازه گیری شد ،دیگر آن کمیت سابق نیست.دو عقیده فوق دارای فلسفه های متفاوتند اما در عمل با هم مطابقت می کنند.آنها دو فرض مختلف را عنوان می کنند اما در عمل به یک نتیجه می رسند.

مشاهداتی که معرف یک کمیت هستند ممکن است همه با هم مساوی و یا دارای اختلافاتی باشند. برای مثال در اندازه گیری ضلع یک مستطیل با خطکش مدرج دو امکان وجود دارد.(شکل 4-1 a,4-1 a)



شکل 4-1

اگر طول ضلع مستطیل دقیقا معادل تعداد عدد صحیح از درجات خط کش باشد ، اندازه گیری های مکرر آن طول سنج ، یک مقدار واحد خواهد شد. برای اینکه ابتدا و انتهای ضلع مستطیل همزمان روبروی درجات معینی از خط کش بوده و تفاضل آن درجات به عنوان طول ضلع در تمام اندازه گیری ها ثابت خواهد ماند اما اگر انتهای ضلع مستطیل بین دو درجه از خطکش واقع شود ، در این صورت ضلع مستطیل برابر تعداد صحیحی از درجات بعلاوه
کسری از یک درجه خط کش خواهد بود که این کسر بوسیله چشم برآورد می شود. این برآورد یا مشاهده در دفعات مختلف اندازه گیری فرق خواهد کرد.این فرق ممکن است به علت مشاهده کننده های مختلف ویا سایر علل باشد.

معمولا وجود اختلاف میان اندازه گیری های یک کمیت فیزیکی دلتیل زیادی دارد از قبیل: طرح اندازه گیری،دستگاه اندازه گیری ،دقت مورد نظر،شرایط جوی وغیره.اگر علت وجود اختلاف ها را بدانیم،میتوانیم به طریقی در رفع آنها اقدام نماییم. یا به عبارت دیگربااعمال تصحیحاتی می توانیم تاثیرات نا خواسته را که معمولا خطاهای سیستماتیک نامیده می شوند از بین برد.خطاهای سیستماتیک متنوع می باشند از قبیل تغییرات طول یک نوار اندازه گیری بر حسب درجه حرارت محیط،تغییرات شرایط جوی ، زمان و غیره.

در عمل می توان خطا های سیستماتیک را از بین برد به شرطی که بتوان تصحیحات مربوطه را به صورت توابع ریاضی از بعضی کمیت های فیزیکی قابل اندازه گیری بیان کرد. در بعضی حالات در طول مشاهدات خطاها هم از نظر قدر مطلق وهم از نظر علامت ثابت می مانند.برای مثال می توان اکثر خطاهای دستگاهی را نام برد.در چنین حالات خطاهای سیستماتیک را با بکار بردن روش های مناسب اندازه گیری از بین می برند.برای مثال خطای قرائت میر ناشی از غیر واقعی بودن محور دیدگانی ترازیاب حین ترازیابی را می توان با مساوی گرفتن فاصله مسیر عقب و جلو از ترازیابی حذف کرد.

در بحث فوق فرض بر این است که اگر اشتباه یا خطاهای بزرگ در مشاهدات وجود ندارد.اشتباه در اندازه گیری ناشی از عدم دقت مشاهده کننده و یا نویسنده می باشد.اشتباهات باید قبل از اینکه کار با مشاهدات شروع شود از بین بروند.رفع اشتباهات از روی مشاهدات طرق مختلف دارد و بستگی به نوع مشاهده دارد که مورد بحث اینجا نیست.

4-2 خطاهای اتفاقی Random(Accidental) Errors

حتی بعد از حذف اشتباهات و اعمال تصحیحات لازم برای حذف خطاهای سیستماتیک هنوز مشاهدات یک کمیت فیزیکی تنها با هم مساوی نبوده و دارای اختلاف کوچکی هستند،این

اختلافات را به گردن یک عده دلایل نا شناخته یا قسمتی ناشناخته انداخت. این اختلافات عملاً اجتناب ناپذیر هستند و می گوییم مشاهدات دارای خطای تصادفی و یا اتفاقی هستند.

برای درک مطالب فوق فرض می کنیم که سری L از مشاهدات یک کمیت فیزیکی ℓ' داده شده است.

$$L = (\ell_1, \ell_2, ..., \ell_n) \quad (4\text{-}1)$$

فرض می کنیم که هر مشاهده $i = 1,2,...,n$ و ℓ_i معرف کمیت ℓ' (ℓ' ارزش واقعی آن کمیت و نا معلوم می باشد) باشد و بتوان نوشت:

$$\ell_i = \ell' + \varepsilon_i \qquad \text{و} \qquad i = 1,2,...,n \quad (4\text{-}2)$$

مقادیر ε_iها را خطا های اتفاقی * می نامند.

* اتفاق می افتد ،در واقع همیشه اتفاق می افتد،که خطاهای ε بستگی به یک یا چند پارامتر نظیر درجه حرارت فشار، زمان وغیره که قبلاً مورد بررسی واقع نشده وحذف نشده اند باشند.در این صورت می گوییم که εها بطور سیستماتیک ویا قابل پیش بینی با پارامترها تغییر می کنند.یا می گوییم یک بستگی بین خطاها و پارامترها وجود دارد یعنی هنوز مشاهدات دارای خطاهای سیستماتیک هستند. در این صورت شاید بتوان با در نظر گرفتن قوانین حاکم بر رفتار آنها خطاهای سیستماتیک را دوباره از بین برد.

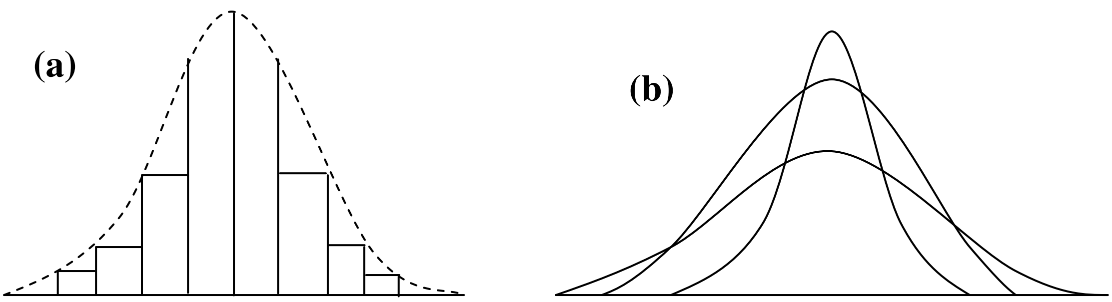
سری

$$\varepsilon = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n) \quad (4\text{-}3)$$

و یا سری L ، معادله 1-4 یک نمونه تصادفی (بخش 1-1-3) نامیده می شود. این نمونه تصادفی دارای یک متغیر تصادفی مادر (2-1-3) خواهد بود.

3-4 تابع PDF گوس (Gauss) ، قانون خطاهای گوس

نمودارها یا پلی گن های نمونه های مشاهدات عموماً بطرف منحنی شکل زنگ متمایل هستند.(شکل a 2-4 و b)



(شکل 2-

اشخاص زیادی در طول تاریخ سعی کرده اند پدیده فوق را تفسیر کنند و یک تئوری برای آن بیان کنند. گوس(Gauss) و لاپلاس (Laplace) دو نفر بودند که بدون اطلاع از کار یکدیگر تفسیر یکسانی از پدیده فوق ارائه دادند. این تفسیر که مورد قبول می باشد منجر به مدل مشهور Gaussian PDF شده است.

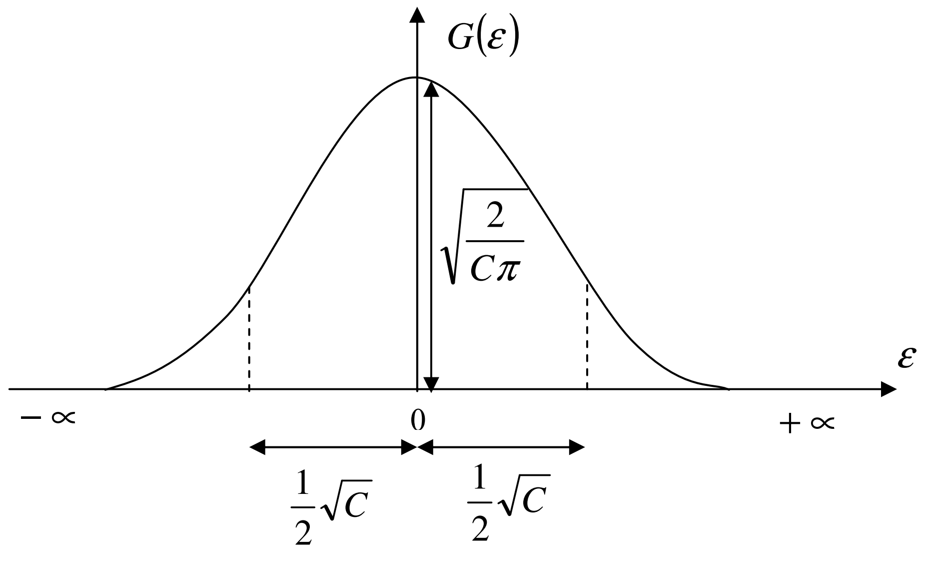
فرضیات لازم، منسوب به Hagen ،که باید مورد نظر قرار گیرد وبدست آوردن قانون فوق ، منسوب به

de Mainer ،در ضمیمه شماره 1 داده شده است.فقط نتیجه بدست آمده را در زیر می نویسیم.

تابع پخش گوس (Gaussian PDF) که با G(c,ε) نشان می دهیم به صورت زیر نوشته می شود:

$$G(c;\varepsilon)=\sqrt{2/c\pi}\,e^{\frac{-2\varepsilon^2}{c}} \qquad (4\text{-}4)$$

در فرمول فوق ε عبارت است از خطای اتفاقی یعنی نوع مخصوصی از متغیر تصادفی که دارای میانگین صفر است و c تنها پارامتر تابع پخش می باشد. تابع پخش گوس یک تابع پیوسته بوده و در شکل 3-4 نشان داده شده است.



(شکل 3-4)

از شکل فوق مشخصات تابع گوس به صورت زیر خواهد بود:

1. تابع گوس نسبت به محور صفر قرینه می باشد.

2. ماکزیمم مقدار تابع در نقطه $\varepsilon=0$ بوده و مساوی $\sqrt{\dfrac{2}{C\pi}}$ می باشد. معلوم است که این مقدار بر حسب پارامتر c تغییر می کند.(شکل 4-2b)

3. مقدار تابع بر حسب مقادیر بزرگ مثبت و یا منفی ε یعنی $\varepsilon\to\pm\infty$ به سمت صفر میل می کند.

4. تابع G دارای دو نقطه عطف با طول های $\varepsilon=\pm\frac{1}{2}C$ می باشد.

شکل تابع G ، قانون مشهور به قانون گوس یک نمونه بزرگ خطا ها را منعکس می کند و مطابق این قانون :

الف- خطاهای کوچک بیشتر از خطا های بزرگ اتفاق (از نظر قدر مطلق) اتفاق می افتند.

ب – خطاهای مثبت و منفی دارای احتمال و قوع یکسان می باشند*.

* نتایج فوق در مورد پخش خطاها می تواند از تئوری حد مرکزی نیز با استفاده از فرضیاتی کمی ضعیفتر (کلی تر) بدست آید.

از آنجایی که تابعG یک تابع پخش احتمال می باشد، شرط زیر در مورد آن برقرار است:

$$\int_{-\infty}^{\infty} G(c;\varepsilon)d\varepsilon=\sqrt{2/c\pi}\int_{-\infty}^{\infty} e^{\frac{-2\varepsilon^2}{c}}d\varepsilon=1 \qquad (4\text{-}5)$$

4-4 میانگین و واریانس تابع گوس (Gaussian PDF)

از آنجایی که تابع G یک تابع قرینه حول محور صفر می باشد،بنابراین میانگین آن برابر صفر است.

واریانس $\sigma_\varepsilon{}^2$ تابع G از فرمول واریانس بدست می آید:

(4-6)

$$\sigma_\varepsilon^2=E^*(\varepsilon-\mu_\varepsilon)^2=\int_{-\infty}^{\infty}\varepsilon^2 G(c;\varepsilon)d\varepsilon=\sqrt{2/c\pi}\int_{-\infty}^{\infty}\varepsilon^2 e^{\frac{-2\varepsilon^2}{c}}d\varepsilon$$

با در نظر گرفتن انتگرال زیر

$$\int_{0}^{\infty} t^2 e^{-a^2t^2}dt=\frac{\sqrt{x}}{4a^3} \quad ,(a>0) \qquad (4\text{-}7)$$

با در نظر گرفتن انتگرال زیر

انتگرال 4-6 به صورت زیر محاسبه می شود

$$\int_{-\infty}^{\infty}\varepsilon^2 e^{\frac{-2\varepsilon^2}{c}}d\varepsilon = 2\int_{0}^{\infty}\varepsilon^2 e^{-a^2\varepsilon^2}d\varepsilon = \frac{\sqrt{\pi}}{2a^3}$$

که در آن $a=\sqrt{\frac{2}{C}}$ می باشد. بنابراین داریم

$$\sigma_\varepsilon^2 = \sqrt{2/c\pi}.\frac{\sqrt{\pi}}{2}(\sqrt{c/2})^3 = \frac{c}{4}$$

و از آنجا مقدار c بر حسب واریانس به صورت زیر در می آید

$$c = 4\sigma_\varepsilon^2 \qquad (4\text{-}8)$$

در نتیجه واریانس ${\sigma_\varepsilon}^2$ و یا انحراف معیار تنها پارامتر تابع G خواهد بود.

با جایگزین کردن مقدار c از فرمول (8-4) در فرمول (4-4) خواهیم داشت

$$G(\sigma_\varepsilon,\varepsilon) = \frac{1}{\sigma_\varepsilon\sqrt{2\pi}}e^{\frac{-\varepsilon^2}{2\sigma_e^2}} \qquad (4\text{-}9)$$

از فرمول 8-4 انحراف معیار σ_ε حساب می شود که برابر است با $\sigma_\varepsilon = \frac{1}{2}\sqrt{C}$ که عبارت است از طول نقطه عطف منحنی G شکل (3-4)

مثال: مطلوب است تعیین مقدار تقریبی احتمال $P(-\sigma_\varepsilon \le \varepsilon \le \sigma_\varepsilon)$، با در نظر گرفتن اینکه ε دارای تابع پخش گوس می باشد.

اول تابع $e^{-\frac{\varepsilon^2}{2{\sigma_\varepsilon}^2}}$ را به سری تیلور بسط می دهیم که محاسبه انتگرال مربوطه امکان پذیر باشد برای این کار تابع e^y را بسط می دهیم :

$$e^{y}=1+y+\frac{y^{2}}{2!}+\frac{y^{3}}{3!}+......$$

$$e^{-\frac{\varepsilon^{2}}{2\sigma_{\varepsilon}^{2}}}=1-\frac{\varepsilon^{2}}{2\sigma_{\varepsilon}^{2}}+\frac{\varepsilon^{4}}{8\sigma_{\varepsilon}^{4}}-\frac{\varepsilon^{6}}{48\sigma_{\varepsilon}^{6}}+......$$

$$P(-\sigma_{\varepsilon}\le\varepsilon\le\sigma_{\varepsilon})=\int_{-\sigma_{\varepsilon}}^{+\sigma_{\varepsilon}}G(\sigma_{\varepsilon},\varepsilon)d\varepsilon=\frac{1}{\sigma_{\varepsilon}\sqrt{2\pi}}\int_{-\sigma_{\varepsilon}}^{+\sigma_{\varepsilon}}e^{-\frac{\varepsilon^{2}}{2\sigma_{\varepsilon}^{2}}}d\varepsilon$$

$$=\frac{1}{\sigma_{\varepsilon}\sqrt{2\pi}}\left[\int_{-\sigma_{\varepsilon}}^{+\sigma_{\varepsilon}}d\varepsilon-\frac{1}{2\sigma_{\varepsilon}^{2}}\int_{-\sigma_{\varepsilon}}^{\sigma_{\varepsilon}}\varepsilon^{2}d\varepsilon+\frac{1}{8\sigma_{\varepsilon}^{4}}\int_{-\sigma_{\varepsilon}}^{+\sigma_{\varepsilon}}\varepsilon^{4}d\varepsilon-...\right]$$

$$\frac{1}{\sigma_{\varepsilon}\sqrt{2\pi}}\left[2\sigma_{\varepsilon}-\frac{1}{2\sigma_{\varepsilon}^{2}}\cdot\frac{2\sigma_{\varepsilon}^{3}}{3}+\frac{1}{8\sigma_{\varepsilon}^{4}}\cdot\frac{2\sigma_{\varepsilon}^{5}}{5}-\frac{1}{48\sigma_{\varepsilon}^{6}}\cdot2\sigma_{\varepsilon}^{7}+...\right]$$

$$=\frac{2\sigma_{\varepsilon}}{\sigma_{\varepsilon}\sqrt{2\pi}}[1-0.167+0.025-0.003]=\frac{2}{\sqrt{2\pi}}[0.855]=0.683$$

$$\Rightarrow P(-\sigma_{\varepsilon}\le\varepsilon\le\sigma_{\varepsilon})=0.683$$

با روش فوق می توان احتمال های زیر را نیز محاسبه کرد که نتایج بدست آمده عبارتند از

$$P(-2\sigma_{\varepsilon}\le\varepsilon\le2\sigma_{\varepsilon})=0.954$$
$$P(-3\sigma_{\varepsilon}\le\varepsilon\le3\sigma_{\varepsilon})=0.997$$

4-5 PDF گوس عمومی یا نرمال

تابع PDF گوس (معادله 9-4) را می توان تعمیم داد بطوری که دارای یک میانگین اختیاری μ_y باشد. برای این کار متغیر ε تابع گوس را به متغیر y در فرمول 9-4 ، تبدیل می کنیم

$$y=\varepsilon+\mu_{y} \tag{4-10}$$

شکل جدید تابع PDF با متغیر y که آنرا تابع PDF نرمال می نامند به صورت زیر خواهد بود

$$N(\mu_y, \sigma_y, y) = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{(y-\mu_y)^2}{2\sigma_y{}^2}} \tag{4-11}$$

اسم نرمال را بخاطر اعتمادی که به قانون پخش خطاهای گوس وجود دارد برای تابع فوق قرار داده اند. اگر یک سری خطا دارای رفتاری مطابق با قانون گوس باشد، هیستوگرامی هم شکل با PDF نرمال نشان دهنده آنها را خطاهای نرمال می نامند، در غیر این صورت آنها را غیر نرمال یا غیر طبیعی وعجیب می نامند.

تابع PDF نرمال دارای دو پارامتر می باشد: میانگین μ_y و انحراف معیار σ_y.

دسته توابع $G(\sigma_y,\mu_y)$ جزو خانواده توابع $N(\mu_y,\sigma_y,y)$ می باشند. شرط زیر در مورد تابع پخش N نیز باید بر قرار باشد

$$\int_{-\infty}^{+\infty} N(\mu_y, \sigma_y, y) dy = \frac{1}{\sigma_y \sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(y-\pi_y)^2}{2\sigma_y{}^2}} dy = 1$$

فرمول تابع CDF مربوط به PDF نرمال به صورت زیر می باشد

$$\psi_N(y) = \frac{1}{\sigma_y \sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{(x-\mu_y)^2}{2\sigma_y{}^2}} dx \tag{4-12}$$

جایی که x متغیر زیر انتگرال می باشد

برای تابع PDF عمومی (نرمال) می توان احتمال های زیر را محاسبه کرد

$$P(\mu_y - \sigma_y \le y \le \mu_y + \sigma_y) \cong 0.683$$
$$P(\mu_y - 2\sigma_y \le y \le \mu_y + 2\sigma_y) \cong 0.954$$
$$P(\mu_y - 3\sigma_y \le y \le \mu_y + 3\sigma_y) \cong 0.997$$

نتایج فوق را با نتایج بدست آمده برای تابع PDF مثلثی شکل در مثال 3-18 مقایسه کنید

4-6 PDF نرمال استاندارد

اگر x یک متغیر نرمال ، یعنی با تابع پخش نرمال ، با میانگین μ_x و واریانس σ^2_x باشد متغیر t را که از ترانسفورماسیون خطی زیر بدست می آید

$$t = \frac{x - \mu_x}{\sigma_x}$$

متغیر نرمال استاندارد می نامند. واضح است که متغیر جدید t دارای میانگین صفر و واریانس واحد خواهد بود. در نظر داشته باشید که روش استاندارد کردن فوق هیچ ارتباطی به تابع پخش x ندارد

از ترانسفورماسیون متغیر نرمال y ، در معادله 4-10 به متغیر نرمال استاندارد t

$$t = \frac{y - \mu_y}{\sigma_y}$$

تابع جدیدی حاصل می شود که آنرا PDF نرمال استاندارد می نامند.

$$N(\mu_t, \sigma_t; t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} = N(0,1,t) = N(t) \qquad (4\text{-}14)$$

میانگین μ_t ، تابع فوق صفر و واریانس آن σ^2_t ، برابر واحد می باشد. تابع فوق یک عضو مخصوص از خانواده توابع پخش نرمال می باشد.

از آنجایی که پارامترهای تابع PDF نرمال استاندارد معین و برابر $\mu_t=0$ و $\sigma_t=1$ می باشند و تابع دارای فقط یک متغیر t می باشد، می توان آنرا به ازاء t های مختلف محاسبه کرده و جدولی از مقادیر آن تنظیم کرد.ضمیمه II-A چنین جدولی را نشان می دهد. این جدول مقادیر PDF نرمال استاندارد را به ازاء مقادیر مختلف از متغیر نرمال استاندارد t نشان می دهد.ملاحظه کنید که شرط سوم یک تابع پخش احتمال در مورد PDF نرمال استاندارد نیز برقرار است

$$\int_{-\infty}^{+\infty} N(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = 1$$

تابع CDF منوطه به N(t) به صورت زیر داده می شود

$$\psi_N(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx \qquad (4\text{-}15)$$

یا

$$\psi_N(t) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_{0}^{t} e^{-\frac{x^2}{2}} dx \qquad (4\text{-}16)$$

جاییکه x متغیر داخل انتگرال می باشد. تابع $\psi_N(t)$ نیز بر حسب مقادیر مختلف t محاسبه و در جدولی تنظیم شده است. این جدول در محاسبات احتمال مورد استفاده قرار می گیرد.

ضمیمه II-B جدولی را که بر اساس فرمول 15-4 و برای t های مثبت* تنظیم شده است نشان می دهد.

* برای مقادیر منفی t نظیر $-t_0$ مقدار احتمال تجمعی $P(t<-t_0) = \Psi_N(-t_0)$ از $\Psi_N(t_0)$ به صورت

$\Psi_N(-t_0) = 1-\Psi_N(t_0)$ محاسبه می گردد.

جدول مذکور سطح زیر منحنی PDF نرمال استاندارد را از ∞- تا نقطه t، $\Psi_N(t)$ ،نشان می دهد.ضمیمه II-C شامل جدول مشابهی است که مقدار ترم دوم معادله 10-4 را بر حسب t های مختلف نشان می دهد. در استفاده از جدول های مختلف فوق برای محاسبه احتمال باید دقت لازم بعمل بیاید تا هر جدول در جای مناسب خود استفاده شود.

ترم دوم در معادله 16-4 به $\frac{1}{2\sqrt{2}}erf(t)$ مشهود است یعنی داریم

$$\psi_N(t) = \frac{1}{2} + \frac{1}{2\sqrt{2}}erf(t) \qquad (4\text{-}17)$$

جاییکه erf(t) را تابع خطا (Error Function) می نامند که دارای شکل زیر است

$$erf(t) = \frac{1}{\sqrt{2\pi}}\int_0^t e^{-\frac{x^2}{2}}dx \qquad (4\text{-}18)$$

جدول نمایش تابع erf(t) به ازاء t های مختلف موجود است*

برای اینکه بتوان از جداول PDF و CDF نرمال استاندارد در محاسبه احتمال های مربوط به یک متغیر تصادفی نرمال با میانگین μ_x و واریانس σ_x^2 استفاده نمود، اول متغیر x را استاندارد می کنیم یعنی آنرا به متغیر t طبق معادله 13-4 تبدیل و سپس t را در جدول مربوطه وارد می کنیم. به طور مثال احتمال $P(x<x_0)$ از رابطه زیر بدست می آید

$$P(x \le x_0) = P(\frac{x-\mu_x}{\sigma_x} \le \frac{x_0-\mu_x}{\sigma_x}) \qquad (4\text{-}19)$$

* تابع erf(t) ، در اغلب زبانهای کامپیوتر، به صورت ساب روتین نوشته شده و در حافظه ماشین وجود دارد. در مواقع لازم می توان ساب روتین را از حافظه ماشین پیدا کرده و تابع erf(t) را به طور دقیق برای هر t دلخواه محاسبه نمود.

طرف دوم رابطه فوق برابر با احتمال $P(t<t_0)$ است که از جدول مربوطه استخراج می شود.

مثال 2-4

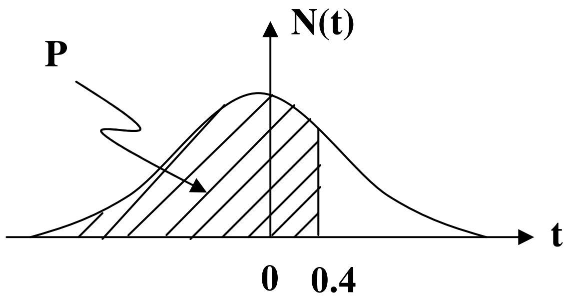
فرض می کنیم که قد دانشجو یک متغیر تصادفی با نام h و با تابع پخش نرمال بوده و دارای میانگین $\mu_h = 66$ و انحراف معیار $\sigma_h = 5$ اینچ باشد. تعداد دانشجویانی را که قد آنها در نامساوی های زیر صدق می کند به صورت عدد k در هزار پیدا کنید.

الف - $h \le 68$ (شکل 4-4a)

ب - $h \le 61$ (شکل 4-4b)

پ - $h \ge 74.6$ (شکل 4-4c)

ت - $[64.3 \le h \le 70]$ (شکل 4-4d)

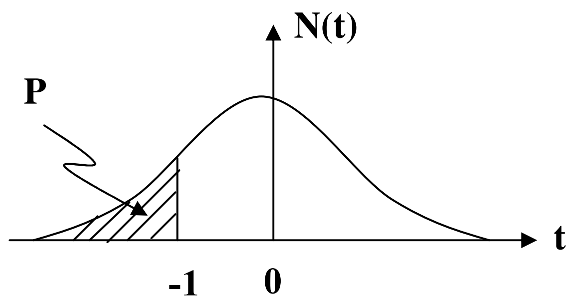


(شکل 4-4a)

الف- در این قسمت از جدول II-B استفاده می کنیم

$$P(h \le 68) = P(t \le \frac{68-66}{5}) = P(t \le 0.4) = 0.6554$$
$$k_1 = (0.6554)(1000) = 655$$

ب-
$$P(h \le 61) = P(t \le \frac{61-66}{5}) = P(t \le -1) = 1 - P(t \le 1) = 1 - 0.8413$$
$$= 0.1587, k_2 = (0.1587)(1000) = 159$$

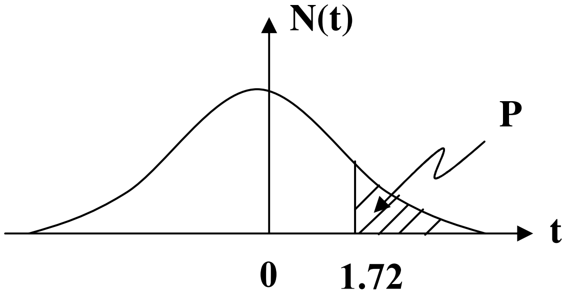


(شکل 4-4b)

پ-

$$P(h \geq 74.6) = P(t \geq \frac{74.6-66}{5}) = P(t \geq 1.72) = 1 - P(t \leq 1.72) = 1 - 0.9573 = 0.0427$$

$$k_3 = (0.0427)(1000) = 43$$



(شکل 4-4c)

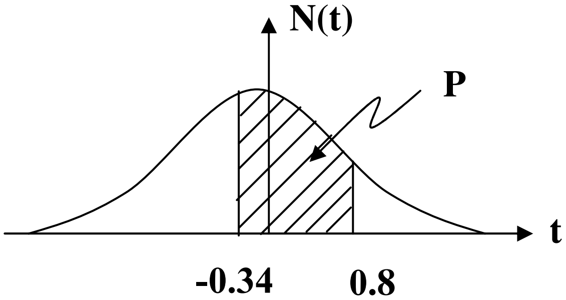
ت-

$$P(64.3 \leq h \leq 70) = P(\frac{64.3-66}{5} \leq t \leq \frac{70-66}{5}) = P(-0.34 \leq t \leq 0.8)$$

$$= P(t \leq 0.8) - P(t \leq -0.34) = P(t \leq 0.8) - \left[1 - P(t \leq 0.34)\right]$$

$$0.7881 - (1 - 0.6321) = 0.7881 - 0.3669 = 0.4212$$

$$k_4 = (0.4212)(1000) = 421$$



(شکل 4-4d)

مثال 4-3

برای متغیر تصادفی نرمال h (مثال 4-2)، قد H دانشجویان را که در نامساوی های زیر صدق می کند را پیدا کنید.

الف- $P(h \leq H_1) = 0.6554$ (شکل 4-5a)

ب - $P(h \geq H_2) = 0.2500$ (شکل 4-5b)

پ - $P(h \leq H_3) = 0.2000$ (شکل 4-5c)

ت - $P(H_4 \leq h \leq H_5) = 0.9500$ جاییکه $H_4 = \mu_h - k$ و $H_5 = \mu_h + k$ (شکل 4-5d)

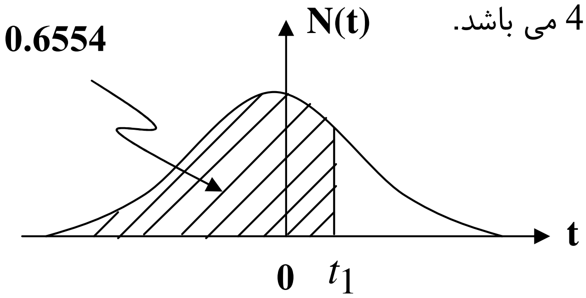
در این مثال از منحنی CDF نرمال استاندارد(ضمیمه II-B) استفاده می کنیم

الف- $P(h \leq H_1) = P(t \leq t_1) = 0.6554$

از جدول مذکور بدست می آید $t_1=0.4$ که مطابق است با احتمال $P=0.6554$. می دانیم که $h_1=\dfrac{H_1-\mu_h}{\sigma_h}$ در مثال 2-4 داریم $\mu_h=66$ و $\sigma_h=5$. بنابراین خواهیم داشت

$$t_1=\frac{H_1-66}{5}\Rightarrow H_1=5t_1+66=5\times0.4+66=68$$

عدد بدست آمده فوق معادل قسمت اول مثال 2-4 می باشد. در حقیقت کاری که در اینجا انجام شده است عکس قسمت الف مثال 2-4 می باشد.



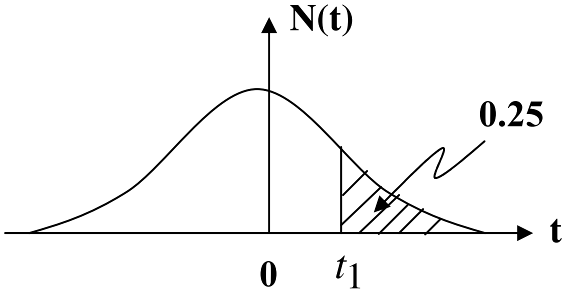
(شکل 4-5a)

ب -

$$P(h\geq H_2)=P(t\geq t_2)=0.25$$
$$P(t\geq t_2)=1-P(t\leq t_2)=0.25, P(t\leq t_2)=1-0.25=0.75$$

با واسطه یابی در جدول فوق الذکر خواهیم داشت $t_2=0.675$ بنابراین داریم

$$t_2=\frac{H_2-\mu_h}{\sigma_h}=\frac{H_2-66}{5}=0.675$$
$$H_2=3.375+66=69.375(inch)$$



(شکل 4-5b)

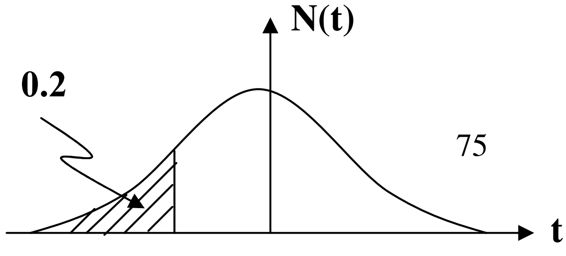
پ - $P(h\leq H_3)=P(t\leq t_3)=0.2$

در جدول فوق الذکر کوچکترین مقداری که برای احتمال مثبت شده عدد 0/5 می باشد. یعنی جدول فوق فقط برای t های مثبت تنظیم شده است. برای اینکه بتوان از آن جدول در اینجا استفاده کرد می نویسیم

$$P(t\leq -t_3)=1-P(t\leq t_3)=1-0.2=0.8$$

دوباره با واسطه یابی خواهیم داشت $-t_3=0.842$ و از آنجا

$$t_3=\frac{H_3-66}{5}=-0.842\Rightarrow H_3=-4.21+66=61.79(inch)$$



(شکل 4-5c)

ت -
$$P(H_4 \le h \le H_5) = P(\frac{H_4-\mu_h}{\sigma_h} \le t \le \frac{H_5-\mu_h}{\sigma_h})5$$
$$= P(\frac{-k}{\sigma_h} \le t \le \frac{k}{\sigma_h}) = P(-t_0 \le t \le t_0) = 0.95$$

که در آن $t_0 = \frac{k}{\sigma_h} = \frac{k}{5}$

$$P(-t_0 \le t \le t_0) = P(t \le -t_0) - P(t \le t_0) = 0.95$$

با در نظر گرفتن حالت تقارن PDF نرمال می توان نوشت

$$P(t \ge t_0) = P(t \le -t_0) = \frac{1-0.95}{2} = 0.025$$
$$P(t \le t_0) = 0.95 + 0.025 = 0.975$$

یا

$$P(t \le t_0) = 1 - 0.025 = 0.975$$

از جدول فوق الذکر بدست می آید $t_0=1.96$ که مربوط به احتمال $P=0.975$ می باشد

$$t_0 = \frac{k}{5} = 1.96 \Rightarrow k = 9.80$$

بنابراین

$$H_4 = \mu_h - k = 66 - 9.8 = 56.2$$
$$H_5 = \mu_h + k = 66 + 9.8 = 75.8$$

مثال 4-4

با استفاده از جدول CDF نرمال استاندارد مثال 1-4 را که در آن احتمال $P(-\sigma_\varepsilon<\varepsilon<\sigma_\varepsilon)$ حساب شده بود را حل می کنیم. ε یک متغیر تصادفی بامیانگین $\mu_\varepsilon=0$ و تابع پخش گوس بود.

$$P(-\sigma_\varepsilon \le \varepsilon \le \sigma_\varepsilon) = P(\frac{-\sigma_\varepsilon-\mu_\varepsilon}{\sigma_\varepsilon} \le t \le \frac{\sigma_\varepsilon-\mu_\varepsilon}{\sigma_\varepsilon})$$
$$= P(-\frac{\sigma_\varepsilon}{\sigma_\varepsilon} \le t \le \frac{\sigma_\varepsilon}{\sigma_\varepsilon}) = P(-1 \le t \le 1)$$

با توجه به شکل 6-4 و در نظر گرفتن حالت قرینه PDF نرمال داریم

$P(-1 \le t \le 1) = 2P(0 \le t \le 1)$

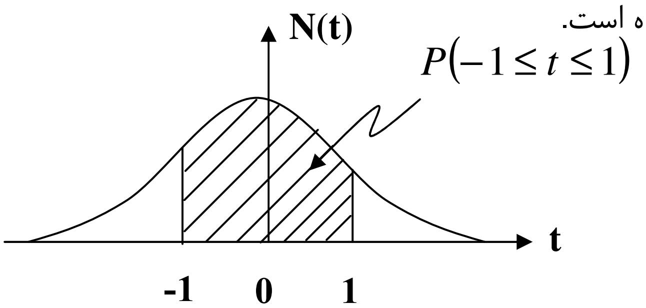
از جدول موجود در ضمیمه II-C بدست می آید

$P(0 \le t \le 1) = 0.3413$

بنابراین داریم

$P(-\sigma_{\varepsilon} \le t \le \sigma_{\varepsilon}) = 2(0.3413) = 0.6826 = 0.683$

نتیجه همان است که در مثال 1-4 بدست آمده است.



(شکل 6-4)

7-4 فرض اساسی تئوری خطاها ، تست

در بخش 2-4 بحث روی نمونه تصادفی مشاهدات ما را نیمه تمام گذاشته و به بررسی و توسعه فرمولهای توابع PDF مورد استفاده در تئوری خطاها پرداختیم. حال برگشته به آنجا وفرض اساسی تئوری خطاها را بیان می کنیم. یک سرس حدود از مشاهدات ما را که معرف یک کمیت فیزیکی می باشد یک نمونه تصادفی قلمداد می کنند. این نمونه تصادفی دارای متغیر تصادفی مادر با تابع پخش نرمال $N(\mu_l,\sigma_l,l)$ می باشد. سایر توابع پخش به ندرت مورد استفاده قرار می گیرد.صحت و ارزش فرض فوق ممکن است مورد آزمایش یا تست قرار گیرد که موضوع بحث ما در اینجا نیست.

گفته می شود که میانگین M_L نمونه L ، میانگینμ_l تابع PDF مادر را تقریب (در اینجا بیشتر کلمه برآورد بجای تقریب بکار خواهد رفت) می کند. همینطور گفته می شود که واریانس S^2_L نمونه L ، واریانس σ^2_l تابع PDF مادر را برآورد می کند.

با در نظر گرفتن نمونه اصلی L

$$L = (\ell_i) = (\ell' + \varepsilon_i), i = 1,2,...., m$$

خواهیم داشت

$$M_L = \frac{1}{n}\sum_{i=1}^{n}\ell_i = \frac{1}{n}\sum_{i=1}^{n}(\ell' + \varepsilon_i) = \frac{1}{n}(n\ell') + \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i = \ell' + M_{\varepsilon} \qquad (4\text{-}20)$$

بنا به فرض گذشته نمونه خطاهای ε دارای تابع پخش مادر $N(0,\sigma_\varepsilon,\varepsilon)$ یعنی $\mu_\varepsilon=0$ است پس باید انتظار داشت که میانگین M_ε به سمت صفر میل کند. یعنی $M\varepsilon \to 0$ ، در این صورت معادله فوق را می توان چنین نوشت

$$M_L \cong \ell' \cong \mu_\ell \qquad (4\text{-}21)$$

باید در نظر داشت که منظور از مقدار نامعلوم ℓ' همان میانگین نامعلوم μ_l تابع پخش مادر ℓ می باشد. می گوئیم که میانگین M_L نمونه L ، مقدار میانگین μ_ℓ را تقریب(برآورد)می کند.
واریانس نمونه L را چنین بسط می دهیم

(4-22)

$$S_L{}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\ell_i - M_L\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\ell_i - \left(\ell' + M_\varepsilon\right)\right]^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i - M_\varepsilon\right)^2 = S_\varepsilon{}^2$$

نتیجه فوق نشان می دهد که واریانس نمونه، S^2_L ، برابر است با واریانس نمونه خطاهای مربوطه ، S^2_ε. به این جهت است که بعضی اوقات واریانس S^2_ε را خطای کوادراتیک متوسط(Mean Squre Error) نمونه L می نامند و با علامت اختصاری MSE نشان می دهند. S_L را جذر خطای کوادراتیک متوسط (Root Mean Squre Error) نامیده و با علامت اختصاری RMS نشان می دهند. بنا به فرض اساسی تئوری خطاها معادله (4-22) به صورت زیر نوشته می شود.

$$S_L{}^2 = S_\varepsilon{}^2 \cong \sigma_\ell{}^2 = \sigma_\varepsilon{}^2 \qquad (4\text{-}23)$$

رابطه فوق می رساند که S^2_L ، واریانس $\sigma_\ell{}^2$ تابع PDF مادر $\varepsilon = \left(\varepsilon_1, \varepsilon_2 ..., \varepsilon_n\right)$ را برآورد می کند.

مثال 5-4

فرض می کنیم که نمونه $L = (2,7,6,4,2,7,4,8,6,4)$ دارای تابع پخش نرمال می باشد.نمونه فوق را به نمونه ای تبدیل کنید که دارای

الف- پخش گوس باشد.

ب - پخش نرمال استاندارد باشد.

نخست میانگین و واریانس نمونه را محاسبه می کنیم.

$$M_L = \frac{1}{10}\sum_{i=1}^{10}\ell_i = \frac{1}{10}(50) = 5$$

$$S_L{}^2 = \frac{1}{10}\sum_{i=1}^{10}\left(\ell_i - 5\right)^2 = \frac{1}{10}(40) = 4$$

بنا به فرض اساسی تئوری خطاها می توان گفت

$$\mu_\ell = M_L = 5 \qquad \text{و} \qquad \sigma_\ell = S_L = 2$$

که در آن μ_ℓ و σ_ℓ به ترتیب میانگین و انحراف معیار تابع PDF نرمال مادر $N(\mu_\ell,\sigma_\ell;\ell)$ می باشد. پارامترهای μ_ℓ و σ_ℓ در تبدیل های فوق مورد استفاده قرار می گیرند.

الف- تابع پخش گوس $G(\sigma_\varepsilon,\varepsilon)$ در اینجا دارای $\sigma_\varepsilon=\sigma_\ell=2$ و متغیر ε که از معادله 4-10 محاسبه می شود می باشد.

$$\varepsilon_i=\ell_i-\mu_\ell=\ell_i-5 \qquad و \qquad i=1,2,\ldots,10$$

بنابراین نمونه جدید که دارای پخش گوس بوده و از نمونه L بدست آمده است عبارت است از

$$\varepsilon=(-3,2,1,-1,-3,2,-1,3,1,-1)$$

ب – پخش نرمال استاندارد $N(t)$ در این مسئله دارای متغیر t است که از رابطه زیر محاسبه می شود

$$t_i=\frac{\ell_i-\mu_\ell}{\sigma_\ell}=\frac{\ell_i-5}{2} \qquad و \qquad i=1,2,\ldots,10$$

بنابراین نمونه جدید که دارای پخش $N(t)$ بوده و از نمونه L بدست آمده است عبارت است از

$$T=(-1.5,1,0.5,-0.5,-1.5,1,-0.5,1.5,0.5,-0.5)$$

4-8 باقیمانده ها، تصحیحات و اختلافات

بطوریکه دیده شد ما نمی توانیم مقدار ℓ' یا μ_ℓ را محاسبه کنیم.تنها کاری که می کنیم عبارت است از محاسبه یک برآورد از آن یعنی $\overline{\ell}$ می باشد

$$\overline{\ell}=M_L=\ell'+M_\varepsilon=\ell'+\overline{\varepsilon} \quad * \qquad (4\text{-}24)$$

با امید اینکه بنابرفرض اساسی تئوری خطاها مقدار $\overline{\varepsilon}$ واقعاً به سمت صفر میل خواهد کرد.

* از حالا به بعد ما علامت $\overline{\ell}$ را برای میانگین نمونه بکار خواهیم برد. علامت (-) بالای ℓ نشان دهنده میانگین یا متوسط ℓ خواهد بود.

تعریف باقیمانده : اختلاف بین مشاهده ℓ_i و میانگین نمونه $\overline{\ell}$ را باقیمانده r_i می نامند

$$r_i=\ell_i-\overline{\ell}=(\ell'+\varepsilon_i)-(\ell'+\overline{\varepsilon})=\varepsilon_i-\overline{\varepsilon} \qquad (4\text{-}25)$$

تعریف تصحیح : باقیمانده ها با علامت عکس تصحیحات می نامند.

باید در نظر گرفت که یک باقیمانده، به طوریکه در بالا تعریف شده است، عبارت است از یک مقدار تعیین شده نه یک متغیر. برای اینکه یک نمونه بخصوص مقدار مشاهده شده ℓ_i و میانگین $\overline{\ell}$ ثابت هستند. به عبارت دیگر، برای اینکه نمونه داده شده، باقیمانده ها فقط به یک طریق محاسبه می شوند. ملاحظه کنید که اختلافات $(\ell_i-\overline{\ell})=r_i$ را باقیمانده ها می نامند نه خطاها. برای اینکه خطاها مطابق $\varepsilon_i=(\ell_i-\mu_\ell)$ تعریف شده اند که در آن μ_ℓ ممکن است مساوی $\overline{\ell}$ نباشد.

درعمل اغلب شنیده می شود که می گویند باقیمانده های مینیمم شده و یا باقیمانده های متغیر که به طور جدی عبارات صحیحی نیستند. اگر کسی بخواهد که باقیمانده ها را به عنوان متغیر ها در نظر بگیرد، مسئله را باید به طور دیگری مطرح کرد.

تعریف اختلاف : تفاوت V_i بین مقدار مشاهده شده ℓ_i و یک مقدار اختیاری ℓ^0 را اختلاف می نامند.

$$V_i = \ell_i - \ell^0 \qquad \text{و} \qquad i = 1,2,\ldots,n \tag{4-26}$$

این اختلافات تابع خطی از ℓ^0 هستند و مقدار آنها با انتخاب ℓ^0 تغییر می کند.

در اینجا می توان از مینیمم کردن اختلاف ها و یا تغییرات اختلاف ها صحبت کرد. اغلب در عمل باقیمانده ها و اختلاف ها با هم اشتباه می شوند.

بجاست حالا یک زوج دیگر از فرمول ها برای محاسبه میانگین $\overline{\ell}$ و واریانس S^2_L نمونه بیان کنیم. این فرمول ها مناسب برای محاسبه نمونه های بزرگ خواهند بود. طریق بدست آوردن این فرمول ها نظیر فرمول های 20-4 و 22-4 و 25-4 و 26-4 می باشد. در زیر فقط خود فرمول ها نوشته شده و طریق بدست آوردن آنها به عهده دانشجویان است.

$$\boxed{\overline{\ell} = \ell^0 + \overline{V}} \tag{4-27}$$

$$\boxed{S_L{}^2 = \frac{1}{n}\sum_{i=1}^{n} r_i^2} \tag{4-28}$$

ℓ^0عبارت است از یک مقدار اختیاری معمولا نزدیک به $\overline{\ell}$.

$$\overline{V} = \frac{1}{n}\sum_{i=1}^{n} V_i \qquad \text{و} \qquad V_i = \ell_i - \ell^0 \tag{4-29}$$

$$r_i = \ell_i - \overline{\ell} \tag{4-30}$$

مثال 6-4

ستون دوم جدول صفحه بعد عبارت است از یک نمونه حاوی ده مشاهده از یک طول میانگین و واریانس نمونه را با استفاده از فرمول های ساده شده در بخش فوق بدست آورید.

برای اینکار $\ell^0 = 972.0$ متر انتخاب می کنیم. در این صورت داریم

$$\overline{V} = \frac{1}{10}\sum_{i=1}^{10} V_i = \frac{1}{10}(10.50) = 1.05 \text{ متر}$$

$$\overline{\ell} = \ell^0 + \overline{V} = 972.0 + 1.05 = 973.05 \text{ متر}$$

$$MSE = S_L{}^2 = \frac{1}{10}\sum_{i=1}^{10} V_i{}^2 = \frac{1}{10}(0.5730) = 0.0573 \text{ مترمربع}$$

$$RMS = S_L \cong 0.24 \text{ متر}$$

یکی از راههای کنترل محاسبات عبارتست از تشکیل مجموع باقیمانده ها که باید صفر باشد یعنی $\sum_{i=1}^{n} r_i = 0$ ، چهارمین ستون از جدول زیر را تماشا کنید.

شماره	ℓ_i متر	$V_i = \ell_i - \ell^0$ متر	$r_i = \ell_i - \overline{\ell} = V_i - \overline{V}$ مترمربع	$r_i{}^2 \times 10^4$ سانتیمتر مربع
1	972.89	0.89	-0.16	256
2	973.46	1.46	0.41	1681
3	973.04	1.04	-0.01	1
4	972.73	0.73	-0.32	1024
5	972.63	0.63	-0.42	1764
6	973.01	1.01	-0.04	16
7	973.22	1.22	0.17	289
8	973.10	1.10	0.05	25
9	972.30	1.30	0.25	625
10	973.12	1.12	0.07	49
$\sum$		10.50	-0.95 +0.95 = 0.00	5740

4-9 امکانات دیگر برای تابع PDF مفروض

تابع PDF نرمال و وابستگان آن تنها توابع شکل زنگ می باشند که نمی توان آنها را در نظر گرفت. تحت شرایط مختلف اندازه گیری، می توان انواع مختلف و مناسب منحنی های شکل زنگ برای تابع PDF در نظر گرفت. بطور کلی این منحنی ها داری بیش از دو پارامتر خواهند بود. بیشتر بودن تعداد پارامتر ها، انعطاف پذیری منحنی PDF را در مطابقت با نمونه واقعی

افزایش می دهد. از طرف دیگر محاسبات مربوط به چنین منحنی مشکل خواهد بود. در اینجا فقط اشاره می کنیم به اینکه اخیرا کوشش هائی در زمینه طرح یک دسته از منحنی هائی که در وسط نسبت به منحنی های PDF نرمال نوک تیز تر هستند انجام شده است. چنین منحنی ها را «Laplacintic» می نامند. تعداد کمی از دانشمندان معتقدند که اگر نمونه های مشاهدات حالت نوک تیزی در وسط را نشان می دهند. این حالت هنوز مورد تأیید اکثریت دانشمندان نبوده و باید مدتی صبر کرد تا کلام قطعی در این مورد انتشار یابد.

بنابراین هنوز تابع PDF نرمال به قوت و اهمیت خود باقی است و احتمالاً همینطور باقی خواهد ماند، برای اینکه نسبتا دارای شکل ساده و کمترین تعداد ممکن پارامتر ها (میانگین و انحراف معیار) می باشد.

4-10 معیارهای دیگر سنجش انحراف

تاکنون با دو معیار سنجش انحراف نمونه سروکار داشتیم، به نامهای (RMS) جذر خطای کوادراتیک میانگین و عرض نمونه (Ra). علاوه بر RMS و Ra ، معیارهای سنجش انحراف زیر اغلب مورد استفاده قرار می گیرند.

خطای متوسط یا میانگین a_e نمونه L به صورت زیر تعریف می شود

$$a_e = \frac{1}{n}\sum_{i=1}^{n}\left|\ell_i - \overline{\ell}\right| = \frac{1}{n}\sum_{i=1}^{n}\left|r_i\right| \qquad (4\text{-}31)$$

که عبارت است از میانگین قدر مطلق باقیمانده ها .

محتملترین خطای P_e نمونه L عبارت است از خطائی که برای آن احتمال

$$P(|r| < P_e) = P(|r| > P_e) = 0.5 \qquad (4\text{-}32)$$

باشد. یعنی احتمال اینکه باقیمانده ای کوچکتر از P_e یا احتمال اینکه باقیمانده ای بزرگتر از P_e باشد 50 درصد است.

محتملترین خطای یک نمونه تصادفی را می توان از منحنی CDF مربوط به قدر مطلق باقیمانده های آن بدست آورد. از منحنی CDF طول مربوط به مقدار تابع CDF = 0.5 همان محتملترین خطاست.

هر دو خطای متوسط، a_e ، و محتملترین خطا، P_e ، میتوانند در مورد پخش های پیوسته نیز تعریف شوند. برای مثال با در نظر گرفتن تابع PDF نرمال $N(\mu_x,\sigma_x,x)$ می توان نوشت

$$a_e = \int_{-\infty}^{+\infty} |x-\mu_x| \phi(x) dx$$

$$= \frac{1}{\sigma_x \sqrt{2\pi}} \int_{-\infty}^{+\infty} |x-\mu_x| e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} dx \qquad (4\text{-}33)$$

با در نظر گرفتن حالت تقارن منحنی PDF نرمال، برای P_e می توان نوشت

$$P(x \le \mu_x - P_e) = \psi_N(\mu_x - P_e) = \frac{1}{\sigma_x \sqrt{2\pi}} \int_{-\infty}^{\mu_x - P_e} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} dx = 0.25 \qquad (4\text{-}34)$$

$$P(x \le \mu_x + P_e) = \psi_N(\mu_x + P_e) = 0.75 \qquad (4\text{-}35)$$

منظور از تابع Ψ_N همان تابع CDF نرمال می باشد.

می توان نشان داد که رابطه تقریبی زیر بین خطاهای σ_x ، a_e و P_e تابع پخش نرمال $N(\mu_x,\sigma_x,x)$ ، بر قرار است.

$$\sigma_x \cong 1.25 a_e \cong 1.5 P_e$$

یا

$$\sigma_x : a_e : P_e \cong 1.0 : 0.80 : 0.67 \qquad (4\text{-}36)$$

P_e خطای محتمل و $\overline{a} \pm d\alpha$ خطای مطلق نامیده می شود.

خطای نسبی r_e نمونه L را به صورت زیر تعریف می شود

$$\boxed{r_e = \frac{S_L}{\overline{\ell}}} \qquad (4\text{-}37)$$

که عبارت است از نسبت بین RMS نمونه و میانگین نمونه.

در عمل از خطای نسبی برای نشان دادن دقت نتبجه یعنی دقت میانگین نمونه استفاده می کنند. در این صورت خطای نسبی به صورت زیر تعریف خواهد شد.

$$\boxed{\overline{r}_e = \frac{S_{\overline{\ell}}}{\overline{\ell}}} \qquad (4\text{-}38)$$

که در آن $S_{\overline{\ell}}$ عبارت است از انحراف معیار میانگین $\overline{\ell}$ که بعداً در فصل 6 بحث خواهد شد.

اغلب شنیده می شود که در مورد دقت نسبی یک کمیت اندازه گیری شده به صورت مثلا 3ppm (یعنی سه قسمت در میلیون) صحبت می کنند که چیزی نیست جز خطای نسبی. یعنی

خطای نسبی آن اندازه گیری برابر 3×10^{-6} می باشد. باید در نظر گرفت که بر خلاف انواع خطاهای مذکور در فوق، خطای نسبی فاقد بعد اندازه گیری می باشد.

فکر ایجاد فواصل اطمینان در اندازه گیری ها مبتنی بر این فرض است که یک نمونه اندازه گیری L دارای تابع پخش نرمال $N(\bar{\ell}, S_L; \ell)$ می باشد. خیلی مرسوم است که یک نمونه را به وسیله میانگین و انحراف معیار آن به صورت زیر نشان می دهند

$$\boxed{\left[\bar{\ell} \pm S_L\right]}$$

یا

$$\boxed{\left[\bar{\ell} - S_L \le \ell \le \bar{\ell} + S_L\right]} \qquad (4\text{-}39)$$

و آنرا " فاصله اطمینان 0.68 احتمال ℓ " می نامند. عدد 0.68 از محاسبه احتمال $P(\mu_\ell - \sigma_\ell \le \ell \le \mu_\ell + \sigma_\ell)$ از منحنی پخش نرمال بدست آمده است.(بخش 5-4)

معنی فاصله (39-4) اینست که 68 درصد احتمال وجود دارد که مقدار واقعی کمیت ℓ در آن فاصله باشد.

همینطور می توان از فاصله اطمینان 0.95 و فاصله اطمینان 0.99صحبت کرد. عموماً فاصله اطمینان برای L را به صورت

$$\left[\bar{\ell} \pm KS_L\right]$$

نشان می دهند و K را چنان تعیین می کنند که احتمال $P(\mu_\ell - K\sigma_\ell \le \ell \le \mu_\ell + K\sigma_\ell)$ برابر 0.95 یا 0.99 باشد. مقادیر $\left[\bar{\ell} + K\sigma_L\right]$ و $\left[\bar{\ell} - K\sigma_L\right]$را به ترتیب حد بالا و حد پایین فاصله اطمینان می نامند.

مثال 7-4

خطای متوسط، خطای نسبی و فاصله اطمینان 0.95 را برای نمونه L در مثال 6-4 محاسبه کنید.

خطای متوسط را با استفاده از فرمول 31-4 و ستون چهارم جدول مثال 6-4 بدست می آورید

$$a_e = \frac{1}{10}\sum_{i=1}^{10}|r_i| = \frac{1}{n}(1.90) = 0.19 \text{ متر}$$

خطای نسبی را از فرمول 37-4 و با استفاده از نتایج بدست آمده در مثال 6-4 محاسبه می کنیم

$$r_e = \frac{S_L}{\bar{\ell}} = \frac{0.24}{973.05} \cong 247\, ppm$$

برای محاسبه فاصله اطمینان 0.95، ضریب k را در فاصله

$$\left[\bar{\ell} - KS_L \le \ell \le \bar{\ell} + KS_L\right]$$

چنان تعیین می کنیم که احتمال

$$P\left(\bar{\ell} - KS_L \le \ell \le \bar{\ell} + KS_L\right) = 0.95$$

برقرار باشد.

احتمال فوق (P(-k<t<k می باشد که از جدول نرمال استاندارد بدست می آید.(مثال 3-4 را ملاحظه کنید). بنابراین خواهیم داشت

$$P(-K \le t \le K) = P(t \le K) - P(t \le -K) = 0.95$$

ازآنجا داریم

$$P(t \le K) = 0.975$$

با استفاده از جدول II-B مقدار k بدست می آید.

$$k=1.96$$

(در عمل k=2 برای فاصله اطمینان 0.95 مورد استفاده قرار می گیرد.)

در نتیجه فاصله اطمینان 0.95 برای کمیت ℓ به صورت زیر خواهد آمد

$$\left[973.05 - 1.96(0.24) \le \ell \le 973.05 + 1.96(0.24)\right]$$

یا

$$\left[972.58 \le \ell \le 973.52\right]$$

یا به عبارت دیگر

$$\left[973.05 \pm 0.47\right]$$ متر

مثال 8-4

متغیر تصادفی x با پخش نرمال N(35,4,x) مفروض است. محتملترین خطای آنرا با استفاده از PDF مفروض بدست آورید.

از تابع PDF مفروض داریم $\mu_x=35$ و $\sigma_x=4$

محتملترین خطای P_e ، خطائی است که در معادله زیر صدق کند.

$$P\left(\mu_x - P_e \le x \le \mu_x + P_e\right) = P\left(-t_P \le t \le t_P\right) = 0.50$$

×××××صفحه های 108 و109 مفقود شده اند×××××

6-فرض می کنیم که نمونه

$$H=(-5,-4,-3,-2,-1,0,1,2,3,4,5)$$

دارای پخش گوس باشد. این نمونه را به یک نمونه جدید چنان تبدیل کنید که

الف- دارای پخش نرمال با میانگین 10 باشد

ب – دارای پخش نرمال استاندارد باشد.

7- متغیر تصادفی x با تابع پخش $N(25,10,x)$ داده شده است. احتمال های زیر را تعیین کنید.

1- $P(x\leq 28.5)$

2- $P(x\geq 27.5)$

3- $P(|x-25|\leq 1.25)$

4- $P(x\leq 22.5)$

5- $P(16.75\leq x\leq 23.82)$

8- متغیر تصادفی x تمرین فوق را در نظر می گیریم. مقادیر z_i را که در معادلات زیر صدق می کنند پیدا کنید.

1- $P(x\leq z_1)=0.65$

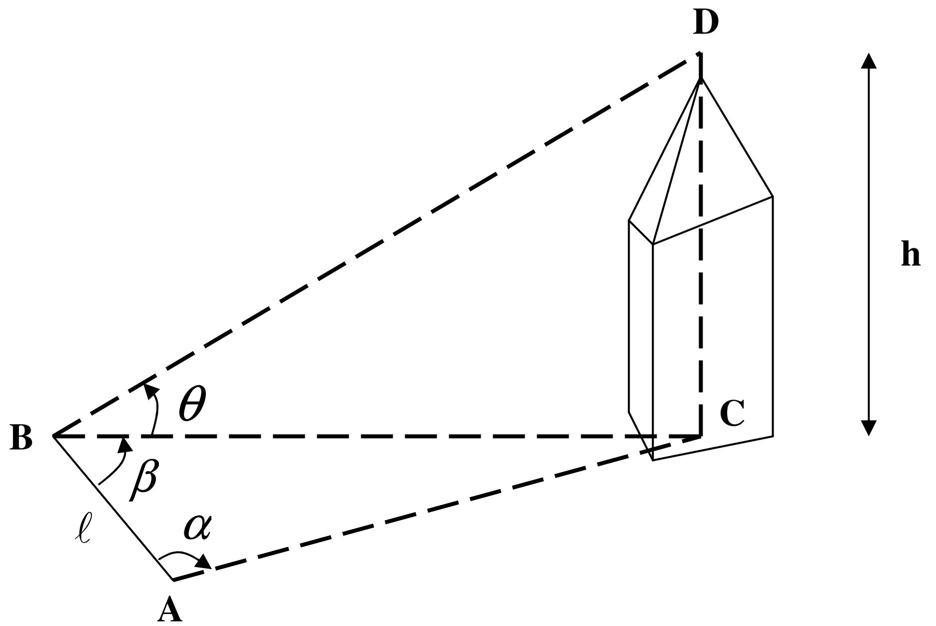
2- $P(x\geq z_2)=0.25$

3- $P(x\leq z_3)=0.33$

4- $P(|x-25|\leq z_4)=0.33$

5- $P(|x-25|\geq z_5)=0.50$

9- شکل زیر روش تعیین ارتفاع برج $CD(h)$ را با اندازه گیری زوایای α وβ وθ و طول ℓ نشان می دهد.



طول افقی AB = ℓ

زوایای افقی در A و B = β و α

زاویه قائم D در B = θ

نتایج اندازه گیری زوایا و طول مذکور (مشاهدات زمینی)در جدول زیر آمده است. درجه حرارت متوسط هوا در موقع اندازه گیری $T=20^{o}$ F بیست درجه فارنهایت بود.اطلاعات زیر در مورد دستگاههای اندازه گیری به نقشه بردار داده شده بود.

الف- میکرومتر دایره قائم تئودولیت، وقتی که حباب تراز مربوطه در وسط قرار می گرفت، بجای عدد "00'00 عدد ("00'03-) را نشان می داد.

ب – دستگاه اندازه گیری طول، یک نوار بیست متری است که در درجه حرارت $T=60^{o}$ F مدرج شده است. ضریب انبساط فلز نوار برابر $\gamma = 5x10^{-5}$ بر یک درجه فارنهایت می باشد.

مشاهدات روی زمین			
ℓ (متر)	α	β	θ
145 . 63	65^{o} $32'$ $03''$	37^{o} $13'$ $08''$	42^{o} $52'$ $15''$
. 55	32 04	13 11	52 30
. 59	31 59	13 10	53 00
. 65	32 01	13 13	51 00
. 58	31 58	13 06	52 15
		13 12	52 45
			51 15
			53 00
			51 45
			52 15

مطلوب است

1. محاسبه بهترین مقدار برای کمیت های ℓ، α ، β و θ .
2. انحراف معیار و خطای متوسط هر کدام از مشاهدات فوق
3. مقایسه دقت مشاهدات فوق(توسط مقایسه خطا های نسبی)
4. با فرض اینکه هر کدام از کمیت ها دارای تابع پخش نرمال است، فاصله اطمینان 0.95 را برای کمیت هاتشکیل بدهید .
5. محاسبه بهترین مقدار برای ارتفاع برج تا نزدیکترین سانتیمتر.

فصل 5 : اصل کمترین مربعات

5-1 میانگین نمونه بعنوان " برآورنده کمترین مربعات "

ممکن است یک نفر این سؤال را مطرح کند : در یک نمونه داده شده L ،
$L=(l_i),i=1,2,\dots,n$ ، مقدار ℓ^0 که مجموع مربعات اختلاف های

$$V_i = \ell_i - \ell^0 \tag{5-1}$$

را کمترین(مینیمم) کند چیست؟

سؤال فوق را می توان دقیق تر به صورت زیر بیان کرد. با تعریف یک "واریانس جدید" بشکل S^{*2}

$$S^{*2} = \frac{1}{n}\sum_{i=1}^{n}\left(\ell_i - \ell^0\right)^2 \tag{5-2}$$

می توان چنین سؤال کرد : مقدار ℓ^0 که تابع S^{*2} را مینیمم می کند چیست؟

بدیهی است که چنین سؤالی را می توان بطور ریاضی جواب داد. از فرمول 5-2 می توان دید که یک تابع از ℓ^0 ، تنها متغیر آزاد در فرمول 5-2 می باشد و می توان نوشت

$$S^{*2} = S^{*2}\left(\ell^0\right) \tag{5-3}$$

می دانیم که تابع S^{*2} به ازاء ℓ^0 وقتی مینیمم است که ℓ^0 جواب معادله زیر باشد

$$\frac{\partial S^{*2}}{\partial \ell^0}=0$$

بنابراین برای پیدا کردن مقدار ℓ^0 کافیست از تابع S^{*2} نسبت به ℓ^0 مشتق گرفته مساوی صفر قرار دهیم

$$\frac{\partial S^{*2}}{\partial \ell^0}=\frac{1}{n}\frac{\partial}{\partial \ell^0}\left[\sum_{i=1}^{n}\left(\ell_i-\ell^0\right)^2\right]=\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\partial}{\partial \ell^0}\left(\ell_i-\ell^0\right)^2\right]$$

$$=\frac{1}{n}\sum_{i=1}^{n}\left[2\left(\ell_i-\ell^0\right)(-1)\right]=-\frac{2}{n}\sum_{i=1}^{n}\left(\ell_i-\ell^0\right)=0$$

و از آنجا داریم :

$$\sum_{i=1}^{n}\left(\ell_i-\ell^0\right)=0$$

از معادله فوق خواهیم داشت :

$$\sum_{i=1}^{n}\ell_i=\sum_{i=1}^{n}\ell^0=m\ell^0$$

و یا

$$\ell^0=\frac{1}{m}\sum_{i=1}^{n}\ell_i\equiv\bar{\ell} \qquad 5\text{-}4$$

نتیجه فوق چیزی نیست جز $\bar{\ell}$ (میانگین نمونه L) . بعبارت دیگر میانگین نمونه مقداری است که مجموع مربعات اختلافها را می نیمم کرده در ضمن اختلافها را تبدیل به باقیمانده ها می کند (بخش 8-4) .

بدلیل فوق میانگین $\bar{\ell}$ نمونه را گاهی اوقات برآورنده کمترین مربعات کمیت ℓ' (μ_ℓ) مینامند . همینطور دیده می شود که مقدار $\bar{\ell}$ می نیمم واریانس را برای نمونه L ایجاد می کند ، اگر واریانس را تابعی از $\bar{\ell}$ در نظر بگیریم .

ملاحظه کنید که خاصیت فوق الذکر میانگین ، هیچ ارتباطی به نوع تابع بخش PDF نمونه ندارد . یعنی که مینگینی $\bar{\ell}$ ، نمونه همیشه یک " برآورنده می فهمیم واریانس "برای نمونه می باشد ، هر شکلی که تابع PDF مربوطه داشته باشد .

5-2) میانگین نمونه بعنوان " برآورنده ماکزیمم احتمال"

نمونه L را در در نظر گرفته و فرض می کنیم که تابع پخش PDF آن شکل نرمال (بخش 5-4) با میانگینی $\mu_\ell = \ell^0$ و واریانس σ_ℓ^2 باشد .

$$\sigma_\ell^2 = S^{-2} = \frac{1}{m}\sum_{i=1}^{n}\left(\ell_i - \ell^0\right)^2 = \frac{1}{m}\sum_{i=1}^{n} u_i^2 \quad (5\text{-}5)$$

می گوییم که تابع PDF نرمال . $N(\mu_\ell, \sigma_\ell; \ell) \equiv N(\ell^0, S^*; \ell)$ مهمترین تابع پخش احتمال برای نمونه $L = (F_i), i = 0,...,r$ می باشد اگر احتمال ؟ وقوع همزمان m عنصر ، که هر کدام دارای تابع پخش $N(\ell^0, S^*; \ell)$ بوده و در همان جائی که L اتفاق می افتد ، ماکزیمم باشد . یعنی می خواهیم که تابع

$$P\left[(\ell_i \le \ell \le \ell_i + \delta\ell_i), i = 1,2,...,n\right] = \prod_{i=1}^{m} N(\ell^0, S^*; \ell_i)\delta\ell_i \qquad (5\text{-}6)$$

نسبت به پارامترهای آزادانه یا متغیرهای موجود ماکزیمم باشد . ℓ^0 تنها پارامتر آزاد تابع 6-5 می باشد . بنابراین احتمال مرکب فوق تابعی است از ℓ^0 .

$$P\left[(\ell_i \le \ell \le \ell_i + \delta\ell_i), i = 1,2,...,n\right] = \lambda(\ell^0) \qquad (5\text{-}7)$$

ملاحظه کنید که مقادیر $\delta\ell_i$ ها بستگی به نمونه L دارند و توسط آن تعیین می شوند . در زیر نشان خواهیم داد که ماکزیمم مقدار تابع $\lambda(\ell^0)$ ، را بر مقداری از ℓ^0 که تابع S^* را می فهمیم می کند حاصل می شود .

$$\max_{\ell^0 \in R}\left[\lambda(\ell^0)\right] = \max_{\ell^0 \in R}\left[\prod_{i=1}^{n} N(\ell^0, S^*; \ell)\delta\ell_i\right]$$

$$= \max_{\ell^0 \in R}\left[\prod_{i=1}^{n}\frac{1}{S^*\sqrt{2\pi}}\prod_{i-1}^{n} e^{-\frac{(\ell_i - \ell^0)^2}{2S^2}}\delta\ell_i\right]$$

$$= \max_{\ell^0 \in R}\left[\left(\frac{1}{S^*\sqrt{2\pi}}\right)^n\prod_{i-1}^{n} e^{-\frac{(\ell_i - \ell^0)^2}{2S^{*2}}}\delta\ell_i\right] \qquad (5\text{-}8)$$

در اینجا $\prod_{i=1}^{n} \delta\ell_i$ توسط نمونه L تعیین می شود و نقشی در عمل ماکزیمم کردن ندارد بنابراین می تواند بعنوان مقدار ثابت در نظر گرفته شود . پس داریم :

$$\max_{\ell^0 \in R} \left[\lambda(\ell^0)\right] = \max_{\ell^0 \in R} \left[\left(\frac{1}{S^* \sqrt{2\pi}}\right)^n \prod_{i-1}^{n} e^{-\frac{(\ell_i - \ell^0)^2}{2S^{*2}}} \delta\ell_i \right] \qquad (5\text{-}9)$$

عبارت دوم سمت راست رابطه فوق را با Q نشان می دهیم .

$$Q = \prod_{i=1}^{n} e^{-x} \qquad \text{و} \qquad x_i = \frac{(\ell_i - \ell^0)^2}{2S^{*2}} \qquad (5\text{-}10)$$

از طرفین رابطه فوق لگاریتم طبیعی می گیریم :

$$\ln Q = \ln\left(\prod_{i=1}^{n} e^{-x_i}\right) = \sum_{i=1}^{n} \ln(e^{-x_i}) = \sum_{i=1}^{n} (-x_i)$$

از آنجا خواهیم داشت :

$$Q = e^{\sum_{i=1}^{n} (-x_i)} \qquad (5\text{-}11)$$

از معادلات 9-5 و 10-5 و 11-5 خواهیم داشت

$$\prod_{i=1}^{n} e^{-\frac{(\ell_i - \ell^0)^2}{2S^{*2}}} = e^{-\frac{1}{2S^{*2}} \sum_{i=1}^{n} (\ell_i - \ell^0)^2} \qquad (5\text{-}12)$$

شرط 9-5 را در مرتبه بصورت زیر می نویسیم .

$$\max_{\ell^0 \in R} \left[\lambda(\ell^0)\right] = \max_{\ell^0 \in R} \left[\left(\frac{1}{S^* \sqrt{2\pi}}\right)^n e^{-\frac{1}{2S^{*2}} \sum_{i=1}^{n} (\ell_i - \ell^0)^2} \right] \qquad (5\text{-}13)$$

از معادله 5-5 داریم

$$\sum_{i=1}^{n} (\ell_i - \ell^0) = nS^{*2}$$

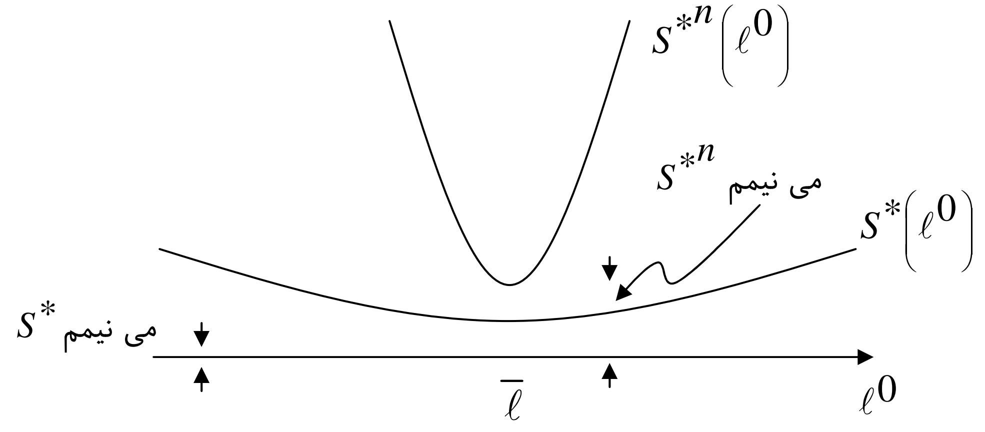
با قرار دادن مقدار فوق در معادله 13-5 خواهیم داشت

$$\max_{\ell^0 \in R} \left\lfloor \lambda(\ell^0) \right\rfloor = \max_{\ell^0 \in R} \left[\left(\frac{1}{S^* \sqrt{2\pi}}\right)^n e^{-\frac{n}{2}} \right] \quad (5\text{-}14)$$

از آنجائیکه در رابطه فوق ، S^* تنها رابطه به ℓ^0 می باشد می توان نوشت :

$$\max_{\ell^0 \in R} \left\lfloor \lambda(\ell^0) \right\rfloor = \max_{\ell^0 \in R} \left[\left(\frac{1}{S^*}\right)^n \right] = \max_{\ell^\circ \in R} \left[S^{*-n} \right] = \max_{\ell^\circ \in R} \left[S^* \right]^n \quad (5\text{-}15)$$

برای اینکه S^* یک تابع غیر منفی از ℓ^0 می باشد . می نیمم مقدار تابع $(S^*)^n$ و می نیمم مقدار تابع S^* با ازاء یک مقدار ℓ^0 حاصل می شود ، شکل 1-5 را ملاحظه کنید .



(شکل 1-5)

سرانجام شرط اصلی (معادله 9-5) را می توان بشکل زیر نوشت .

$$\max_{\ell^0 \in R} \left\lfloor \lambda(\ell^0) \right\rfloor = \max_{\ell^\circ \in R} \left[S^*(\ell^\circ) \right] = \max_{\ell^\circ \in R} \left[S^{*2}(\ell^\circ) \right] \qquad (5\text{-}16)$$

رابطه فوق نشان می دهد که ℓ^0 جواب معادله زیر می باشد .

$$\frac{\partial S^*}{\partial \ell^\circ} = \frac{\partial S^{*2}}{\partial \ell^\circ} = 0$$

یعنی

$$\frac{\partial}{\partial \ell^{\circ}} \sum_{i=1}^{n} V_i^2 = 0 \qquad (5\text{-}17)$$

بدیهی است که شرط فوق همان شرطی است که برای واریانس می نیمم در بخش قبلی داشتیم و نتیجه آن

$$\ell^{\circ} = \overline{\ell}$$

می باشد .

می بینیم که با فرض تابع پخش زمان برای نمونه یک برآورنده ماکزیمم احتمال کمیت ℓ می باشد . در واقع در می یابیم که شرط لازم برای احتمال ماکزیمم عبارتست از

$$\frac{\partial S^{*}}{\partial \ell^{\circ}} = 0 \qquad (5\text{-}18)$$

5-3) اصل کمترین مربعات

دیدیم که میانگین نمونه همیشه باعث می شود که مجموع مربعات اختلافها به می نیمم مقدار خود برسد و تحقق این شرط لازم است تا یک خانواده بزرگ از PDF را برای تابع احتمال همزمان با ماکزیمم احتمال باشد . بنابراین میانگین نمونه در عین حالیکه می نیمم مجموع مربعات اختلافها را ایجاد می کند مهمترین مقدار برای میانگین μ_ℓ نیز است بشرطی که تابع PDF مربوطه قرینه باشد این شرط لازم و کافیست برای میانگین نمونه که هم برآورنده کمترین مربعات و هم برآورنده ماکزیمم احتمال باشد یعنی هر دو برآورنده یکی باشند .

* در تمام بحث هایی که تا به حال انجام گرفت چیزی راجع به مهمترین مقدار σ_ℓ گفته نشد σ_ℓ را باید مطابق فرمول 2-4 فرض کرد .

فکر می نیمم کردن مجموع مربعات اختلافها سرشکنی اصل کمترین مربعات مشهور است و دارای اهمیت اساسی در محاسبات سرشکنی می باشد . بعدا نشان خواهیم داد که چطور اصل کمترین مربعات در برآوردهای مختلف (نه تنها میانگین نمونه) مورد استفاده قرار می گیرد و چطور بصورت متد کمترین مربعات در می آیند . در هر صورت محدودیت های اساسی اصل کمترین مربعات را نباید از نظر دور داشت که عبارتند از :

1- فرض بنا به یک تابع PDF نرمال یا تابع قرینه می باشد .

2- اصل کمترین مربعات چیزی در مورد مهمترین برآونده پارامتر σ_ℓ نسبت به میانگین μ_ℓ تابع PDF فرض شده صحبت نمی کند .

* اگر مقدار بیشتری از خصوصیات نمونه تصادفی معلوم و یا فرض شده باشد می توان بعضی مشخصات انحراف معیار S را نشان داد (بخش 4-2-3) را ملاحظه کنید) .

5-4) اصل کمترین مربعات برای چند متغیری تصادفی

اصل کمترین مربعات نشان داد که میانگین نمونه ، $\bar{\ell}$ و برآورنده میـانگین مـادر μ_ℓ ، تحـت شرایط می نیمم بودن مجموع مربعـات اختلافهـا ، برابرنـد . همینطـور دیـدیم کـه میـانگین $\bar{\ell}$ مهمترین برآورد μ_ℓ می باشد با فرض اینکه متغیر مادر دارای تابع بخش PDF نرمـال و یـا هر تابع قرینه دیگر باشد . حال نشان خواهیم داد که اصل فوق در مورد یک چنـد نمونـه ای بـا فرض اینکه چه متغیری مستقل باشد نیز صادق است .

با نشان دادن چند متغیری با علامت $\tilde{L}$ و مؤلفه های آن با L^j برای j=1,2,…,s و بـا علـم بـــه اینکـــه L^j خـــود یـــک نمونـــه تـــصادفی اســـت مـــی تـــوان نوشـــت (5-19)

$$\tilde{L} = (L^1, L^2, \ldots, L^s)$$

$$L^j = (\ell_1^j, \ell_2^j, \ldots, \ell_n^j) \in R^s$$

فرض می کنیم L_0 یک مقدار بخصوصی برای چند نمونه ای L باشد .

$$L^j = (\ell_0^1, \ell_0^2, \ldots, \ell_0^s) \in R^s \qquad (5\text{-}20)$$

L_0 یک بردار با مؤلفه های عددی ، با یک سری از اعداد حقیقی ، می باشد . اختلاف های مربوطه $\tilde{V}$ که می تواند خود بعنوان یک چند نمونه ای تلقی شود عبارت خواهد بود از

$$\tilde{V} = (\tilde{L} - L_0) \equiv (V^1, V^2, \ldots, V^s) \qquad (5\text{-}21)$$

در اینجا V^j ، بازای j=1,2,…,s ، بنوبه خود یک نمونه از اختلافها می باشد .

$$V^j \equiv (V_1^j, V_2^j, \ldots, V_{n_j}^j) \in R^s \qquad (5\text{-}22)$$

با استفاده از فرمول (52-3) ، می توان فرمولی نظیر فرمول (2-5) برای مؤلفه هـای چندنمونـه ای بصورت زیر نوشت .

$$S_j^{*2} = \frac{1}{n_j}\sum(\ell_i^j - \ell_0^j)^2 = E\left[(V^j)^2\right] \qquad (5\text{-}23)$$

می نیمم کردن واریانس ها یعنی می نیمم کردن هر کدام از $E\left[(V^j)^2\right]$ ها معادل مـی نـیمم کردن هر کدام از S_j^{*2} ها است .

$$\left(\min_{L_0 \in R^s}\left[E(V^j)^2\right], j = 1,2,\ldots,S\right) \equiv \min_{L_0 \in R^s}\left[trace \sum \tilde{L}\ \right] \quad * \qquad (5\text{-}24)$$

که در آن $\Sigma_{\tilde{L}}$ عبارتست از ماتریس واریانس کوواریانس چند نمونه ای $\tilde{L}$ (بخش 6-3-3) با دنبال کردن روشی نظیر بخش 1-5 می توان دید که بردار

$$L_0=(\ell^{-1},\ell^{-2},\ldots,\ell^{-s})\in R^s \qquad (5\text{-}25)$$

شرط 25-5 را بر قرار می کند . از طرف دیگر نتیجه 25-5 چیزی نیست به جز میانگین چند نمونه ای مزبور یهنی

$$L_0=\bar{L}\in R^s \qquad (5\text{-}26)$$

* trace یک ماتریس عبارتست از مجموع عناصر روی قطر اصلی آن

با فرض یک تابع PDF نرمال $N(\ell_0^j,S_j^*,\ell^j)$ برای هر کدام از مؤلفه های L^j چند نمونه ای $\tilde{L}$، تابع PDF مربوطه به چند متغیری مادر را می توان بصورت

$$\phi(\tilde{\ell})=\prod_{j=1}^{s}N(\ell_0^j,S_j^*,\ell^j)=\prod_{j=1}^{s}\frac{1}{s_j^*\sqrt{2\pi}}e^{-\frac{(\ell^j-\ell_0^j)^2}{2S_j^{*2}}} \qquad (5\text{-}27)$$

نوشت که در آن ℓ^j عبارتست از متغیر تصادفی با میانگین ℓ_0^j از انحراف معیار S_j^* با بکار بردن روشی نظیر بخش 2-5 نتیجه می گیریم که بردار

$$L_0=(\ell^{-1},\ell^{-2},\ldots,\ell^{-s})\in R^s \qquad (5\text{-}28)$$

احتمالاً اینرا که اعضای چند متغیری مادر در یکی بعنوان اعضای چند نمونه ای $\tilde{L}$اتفاق بیفتد به ماکزیمم میرساند .

در نتیجه بردار $\bar{L}\in R^S$ ، تحت شرایط فوق * ، برآورنده ، ماکزیمم احتمال میانگینی اگر (میانگین تابع PDF مادر فرض شده) می باشد .

$$\tilde{\mu}=(\mu_1,\mu_2,\ldots,\mu_s)\in R \qquad (5\text{-}29)$$

5-5) تمرین شماره 5

1- ثابت کنید که میانگین μ یک تابع پیوسته PDF - $\phi(x)$ ، بصورت

$$\mu=\int_{-\infty}^{+\infty}x\phi(x)dx$$

واریانس σ^2 تعریف شده بصورت زیر را می نیمم می کند .

$$\sigma^2 = \int_{\infty}^{\infty} (x-\mu)^2 \phi(x) dx$$

2- ثابت کنید که شرط لازم و کافی برای اینکه تابع یکنواخت $R(\ell^\circ, S^*, \ell)$ مهم ترین تابع PDF نمونه L با میانگین $\overline{\ell}$ و واریانس S^2 باشد عبارتست از

$$\frac{\partial S^*}{\partial \ell^\circ} = 0$$

معادله ریاضی تابع یکنواخت R در مثال 3-17 داده شده است .

3- مسئله فوق را در مورد تابع مثلثی $T(\ell^\circ, S^*, \ell)$ که فرم ریاضی آن در مثال 3-18 بخش 3-2-5 داده شده است حل کنید .

* می توان نشان داد که $\overline{L}$ ، برآورنده ، ماکزیمم احتمال میانگین دیگر می باشد حتی اگر یک تابع چند بعدی رابطه (غیر مستقل) PDF برای چند نمونه ای فرض شود .

فصل 6- اساس محاسبات سر شکنی

6-1) نمونه های تصادفی اولیه و بدست آمده

تا کنون با نمونه های تصادفی یک متغیری (یا چند متغیری) سروکار داشتیم که از طریق اندازه گیری و یا نوعی جمع آوری اطلاعات بدست آمده بودند . آنها را نمونه های اولیه و یا اصلی می نامیم .

در عمل ما دنبال نمونه هایی مستقیم که مستقیماً قابل تهیه نیستند بلکه طی محاسباتی از نمونه های اولیه و یا اصلی بدست می آیند . آنها را نمونه های بدست آمده می نامیم .

از نقطه نظر علمی فرق چندانی بین آن دو نوع نمونه وجود ندارد . بطوریکه یک نمونه اولیه را می توان یک نمونه بدست آمده از پدیده های فیزیکی و یا تأثیرات فیزیکی در نظر گرفت . باید هر دو نمونه اولیه و بدست آمده از آن مشخص و معین باشند تا بتوان در مورد تبدیل از یکی به دیگری صحبت کرد .

6-2) انتقال آماری ، مدل ریاضی

عبور از یک نمونه اولیه به نمونه بدست آمده را که تؤام با انتقال واریانس ها و کروواریانس های مربوطه باشد انتقال آماری می نامند . قبلاً دو مثال از این نوع انتقال را در مورد متغیر تصادفی ، نه یک نمونه ، دیدیم (بخش 5-4 و 6-4) اینها انتقال PDFگرمی به PDF نرمال و به PDF نرمال استاندارد بودند .

انتقال آماری همیشه مثل دو انتقال فوق ساده نیست . در واقع ممکن است حتی امکان بدست آوردن نمونه دلخواه از نمونه اولیه بطور کلی وجود نداشته باشد . این حالت معمولاً در مورد چند نمونه ها وجود دارد . یا بعبارت دیگر ممکن است بتوان همیشه نمونه دلخواه (بدست آمده) را بر حسب نمونه اولیه بطور صریح فرموله کرد .

چند نمونه ای اولیه $L \equiv (L^i), i = 1,2,\ldots,s$ را که دارای σ مؤلفه است در نظر می گیریم . هر مؤلفه $L^i \equiv (\ell^i_R), R = 1,2,\ldots,n_i$ خود یک نمونه تصادفی و معرف کمیت فیزیکی ℓ_i می باشند . حال می خواهیم چند نمونه ای X را که دارای n مؤلفه می باشد از چند نمونه ای اولیه فوق بدست می آوریم .

$$X = (X^j), j = 1,2,\ldots,n$$

هر مؤلفهٔ X^j خود یک عدد تصادفی از یک کمیت فیزیکی $X_j, j = 1,2,\ldots,n$ می باشد .

فرمولها یا روابطی که دو کمیت فیزیکی x, ℓ ،

$$\ell = (\ell_1, \ell_2, \ldots, \ell_s) \qquad * \ (6\text{-}1)$$

$$x = (x_1, x_2, \ldots, x_n)$$

را به هم مربوط می سازند ، مدل ریاضی نامیده می شوند و معمولا بصورت زیر نشان داده می شوند.

$$F(\ell, x) = 0 \qquad (6\text{-}2)$$

جائیکه F یک بردار از توابع $f_i, i = 1,2,\ldots,r$ که بین x, ℓ برقرار هستند می بشد . یا بعبارت دیگر F یک بردار با r مؤلفه که هر یک از مؤلفه ها تابعی ، f_i ، است بین ℓ, x .

برای اینکه بتوان x را از ℓ بدست آورد باید مدل ریاضی (6-2) بصورت زیر نوشته شود .

$$x = F(\ell) \qquad (6\text{-}3)$$

که x را بصورت یک تابع صریح از ℓ مشخص می کند .

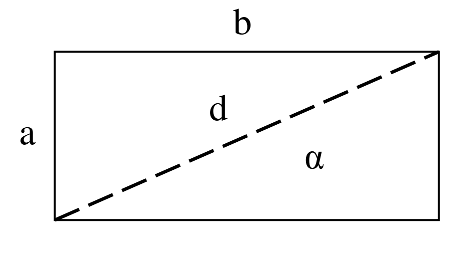
مثال 1-6)

طول b و عرض a یک میز مستطیل شکل ، اندازه گیری شده است . فرض کنید که می خواهیم اطلاعاتی در مورد طول قطر d و مساحت میز α بدست آوریم در اینصورت مدل ریاضی به صورت زیر نوشته می شود .

$$x = F(\ell)$$

$$x = (x_1, x_2) = (d, \alpha)$$

$$\ell = (\ell_1, \ell_2) = (a, b)$$



شکل 1-6

برای بدست آوردن مؤلفه های x از مؤلفه های ℓ روابط زیر را می نویسیم .

$$d = f_1(\ell) = f_1(a,b) = \sqrt{a^2 + b^2}$$
$$\alpha = f_2(\ell) = f_2(a,b) = ab$$

روابط فوق را بصورت بردار می نویسیم .

* x,ℓ چیزی نیستند جز چنر متغیرهای تصادفی مربوط به چند نمونه ای ها ی X , L .

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} d \\ \alpha \end{bmatrix} = \begin{bmatrix} \sqrt{a^2 + b^2} \\ ab \end{bmatrix}$$

هر تابع غیر خطی باید به تابع خطی با استفاده از بسط تیلور و مک لرن تبدیل شود .

امکان انجام انتقال آماری بستگی به سه فاکتور زیر دارد :

1- یعنی امکان فرموله کردن x بطور صریح بر حسب ℓ $x = F(\ell)$

2- کامل بودن چند نمونه ای اولیه L ، بدین معنی که آیا تعداد عناصر نمونه های تشکیل دهنده چند نمونه ای ساده هستند تا بتوان ماتریس واریانس کوواریانس $\sum_L$ را تشکیل داد یا نه.

3- تمایل ما برای امتحان تک تک مؤلفه های چند نمونه ای اولیه L ، تعداد s مؤلفه ، بوسیله n مؤلفه چند نمونه ای X ، که مشکل زیادی را ایجاد می کند .

ارضای دو فاکتور آخری با اندازه ای مشکل است که در مورد آنها حتی اقدام به انتقال آماری نمی کنیم و بجای آن فقط مقادیر $\tilde{E}(x)$, $\sum_X$ برآورد می شوند .

برای اینکار ااول مقادیر $\tilde{E}(L)$, $\sum_L$ از چند نمونه ای اولیه L را تعیین می کنیم . یعنی توسط آنها مقادیر $\tilde{E}(x)$, $\sum_X$ را برآورد می کنیم .

بنابر فرض اساسی تئوری خطاها و برای ایجاد سهولت از محاسبات بعدی ما معمولاً در این مرحله تابعی برای PDF چند متغیری مادر چند نمونه ای L در نظر می گیریم و فرض می کنیم

$$\tilde{E}(L) = \tilde{E}^*(\ell), \sum_L = \sum_L^* \qquad (6\text{-}4)$$
$$\tilde{E}(x) = \tilde{E}^*(x), \sum_x = \sum_x^*$$

درست بهمان طریقی که در حالت یک متغیری فرض کردیم (بخش7-4) یعنی

$$\overline{\ell} = \mu_L, S_L = \sigma_L$$

فرض فوق ما را قادر می سازد که با متغیرهای پیوسته و مدل ریاضی کار کنیم و بنویسیم

$$F(L, X) = 0 \qquad (6\text{-}5)$$

با در نظر گرفتن اینکه هر مقدار X دارای مقدار مربوطه L می باشد .

از حالا به بعد ما علامت $\bar{L}$ را بجای $\tilde{E}(L)$ و علامت $\hat{X}$ را برای نشان دادن برآورد آماری از X بکار خواهیم برد .

در اینصورت معادله 6-5 بشکل زیر در می آید

$$F(\bar{L}, \hat{X}) = 0 \qquad (6\text{-}6)$$

که متشکل از تعدا r معادله بین $\hat{X}, \bar{L}$ می باشد .

از نقطه نظر مدل ریاضی فوق . انتقال آماری در حالت $S \geq n$ ((n تعدا مجهولات و s تعداد مشاهدات)) ممکن و در حالت s<n غیر ممکن خواهد بود . در صورت قابل حل بودن مدل ریاضی فوق ($S \geq n$) تازه امکان دو حالت کاملاً مختلف زیر وجود دارد .

1- مدل دارای یک جواب $\hat{X}$ (r = s = n) ،((n تعداد مجهولات ، s تعداد مشاهدات ، r تعداد معادلات)) است که با استفاده از روش معمول ریاضی بر حسب $\bar{L}$ بدست می آید .

2- مدل دارای اطلات بیشتر از مورد نیاز (r , s > n) است .

در این حالت با استفاده از روش معمولی ریاضی نمی توان جواب واحدی برای مدل بدست آورد. برای اینکه بی نهایت جواب می توان برای مدل بدست آورد .

دو مثال 6-1 ، مدل ریاضی دارای حالت اول می باشد و تعیین $\hat{X}$ از $\bar{L}$ به سادگی انجام می گیرد و تنها مسئله ، محاسبه $\sum_X$ از L و $\sum_L$ می باشد . این مسئله از طریق قانون معروف " انتشار خطاها " که مورد بحث بخش بعدی است حل می شود .

در حالت مدلهای با استفاده بیش از مجهولات (یا مشاهدات زیادتر از مجهولات) انتقال آماری ($(\bar{L}, \sum L) \to (\hat{X}, \sum X)$ ، مسئله سر شکنی را تشکیل می دهد . *

* باید گفت که در عمل در هر دو حالت ما با $\sum_{\bar{L}}, \sum_{\hat{X}}$ واریانس کوواریانس $\hat{X}, \bar{L}$ کار می کنیم نه با $\sum_L, \sum_X$ متعلق به نمونه های L , X ماتریس های $\sum \bar{L}, \sum \hat{X}$ در بخش 6-4-4 محاسبه خواهد شد .

6-3) انتشار خطاها

6-3-1 انتشار ماتریس واریانس کوواریانس ، قانون کوواریانس رابطه بین $\sum_L, \sum_X$ را در مدل ریاضی

$$F(\bar{L}, \hat{X}) = 0$$

انتشار ماتریس واریانس کوواریانس می نامند . چنین رابطه ای را فقط می توان برای یک مدل صریح در $\hat{X}$ ،

$$\hat{X} = F(\bar{L})$$

بدست آورد . برای سهولت اول رابطه مزبور را برای یک مدل صریح خطی ،

$$\hat{X} = B\bar{L}$$

بدست می آوریم . B یک ماتریس دارای m سطر و s ستون از عناصر مشخص و ثابت می باشد *.

* ماتریس B را که رابطه خطی بین X و L را نشان می دهد ، ماتریس ضرایب L در مدل خطی می نامند یا فقط ماتریس ضرایب می نامند.

معادله فوق دارای یک جواب برای $\hat{X}$ ، آن طوری که مورد نیاز است می باشد. ما میخواهیم انتقال زیر را انجام دهیم .

$$\sum_L = \tilde{E}((L-\bar{L})(L-\bar{L})^T) \rightarrow \sum_X \qquad (6\text{-}8)$$

که در آن $\bar{L} = \tilde{E}(L)$ می باشد . از تعریف ماتریس واریانس کوواریانس شروع می کنیم و مینویسیم

$$\sum_X = \tilde{E}((X-\tilde{E}(X))(X-\tilde{E}(X))^T) \qquad (6\text{-}9)$$

در اینجا X =BL بوده و برای $\tilde{E}(X)$ می نویسیم

$$\tilde{E}(X) = \tilde{E}\Big((X-\tilde{E}(X))(X-\tilde{E}(X))^T\Big)$$

$$\sum_X = \tilde{E}((BL-B\bar{L})(BL-B\bar{L})^T)$$

$$= B\tilde{E}((L-\bar{L})(L-\bar{L})^T)B^T = \tilde{E}(B(L-\bar{L})(L-\bar{L})^T B^T)$$

$$= B\sum_L B^T$$

$$\boxed{\Sigma_X = B\Sigma_L B^T} \qquad (6\text{-}10)$$

فرمول فوق به قانون انتشار ماتریس واریانس ـ کوواریانس و یا به طور ساده قانون واریانس معروف است .

مثال 6-2

فرض می کنیم که ماتریس واریانس کوواریانس چند نمونه ای $L = (\ell_1, \ell_2, \ell_3)$ بشکل زیر داده شده است .

$$\Sigma_L = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & 4 \end{bmatrix}$$

اگر چند نمونه ای $X=(x_1,x_2)$ دارای رابطه زیر با چند نمونه ای L باشد

$$x_1=\ell_1-3\ell_3$$
$$x2=2\ell_1+\ell_2$$

ماتریس واریانس ـ کوواریانس $\sum_X$ را پیدا کنید .

دیده می شود که روابط فوق بین مؤلفه های L , X خطی بوده و بشکل ماتریسی زیر می باشد .

$$X_{2,1}=B_{2,3}L_{3,1}$$

یعنی

$$\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}1&0&-3\\2&1&0\end{bmatrix}\begin{bmatrix}\ell_1\\\ell_2\\\ell_3\end{bmatrix}$$

ماتریس $\sum_X$ از معادله 6-10 محاسبه می شود :

$$\Sigma X_{2,2}=B_{2,3}\Sigma L_{3,3}\ B^T_{3,2}$$

$$\sum_X=\begin{bmatrix}1&0&-3\\2&1&0\end{bmatrix}\begin{bmatrix}3&2&0\\2&3&1\\0&1&4\end{bmatrix}\begin{bmatrix}1&2\\0&1\\-3&0\end{bmatrix}$$

$$=\begin{bmatrix}3&-1&-12\\8&7&1\end{bmatrix}\begin{bmatrix}1&2\\0&1\\-3&0\end{bmatrix}=\begin{bmatrix}39&5\\5&23\end{bmatrix}$$

$$\sum_X=\begin{bmatrix}39&5\\5&23\end{bmatrix}$$

قانون کوواریانس یا انتشار ماتریس واریانس کوواریانس را می توان در حالت عمومی مدل ریاضی یعنی برای مدل غیر خطی

$$X=F(L) \qquad (6\text{-}11)$$

نیز بدست آورده جائیکه تابع F دارای حداقل مشتق مرتبه اول می باشد . برای اینکار نخست از تقریب دیگری استفاده کرده مدل را خطی می کنیم . برای خطی کردن مدل 6-11 ، آنرا مثلاَ حول یک نقطه تقریبی L^0 به سری تیلور بسط می دهیم .

$$X=F(L^\circ)+\left.\frac{dF}{dL}\right|_{L=L_0}(L-L^\circ)+$$ نرمالهای با درجه بالاتر

جائیکه

$$\left.\frac{dF}{dL}\right|_{L=L_0}(L-L^\circ)=\sum_{i=1}^{n}\left.\frac{\partial F}{\partial \ell_i}\right|_{\ell_i-\ell_i^\circ}(\ell_i-\ell_i^\circ)$$

دو ترم اول از بسط فوق را فقط برای X در نظر می گیریم . این کار مجاز است برای اینکه عناصر ماتریس $\sum_L$ بمراتب کوچکتر از مقادیر ℓ_i ها بوده و می توان نوشت .

$$X=F(L^\circ)+B(L-L^\circ) \qquad (6\text{-}12)$$

جائیکه B دوباره یک ماتریس n سطر در s ستون را تشکیل از مشتقات جزئی * $\left.\frac{\partial x_i}{\partial \ell_j}\right|_{\ell_i-\ell_i^\circ}$ می باشد . با اعمال اپراتور E روی X در نظر می گیریم . این کار مجاز است برای اینکه عناصر ماتریس $\sum_L$ قانون کوواریانس بدست می آید . باید در نظر گرفت که

$$\tilde{E}(L^\circ)=L^\circ,\tilde{E}(F(L^\circ))=F(L^\circ)$$

$$\tilde{E}(X)=F(L^\circ)+B(\tilde{E}(L)-L^\circ) \qquad (6\text{-}13)$$

$$X-\tilde{E}(X)=B(L-E(L))=B(L-\overline{L}) \qquad (6\text{-}14)$$

$$\begin{aligned}\sum\nolimits_X &=\tilde{E}\Big((X-\overline{E}(X))(X-\tilde{E}(X))^T\Big)\\ &=\tilde{E}\Big(B(L-\overline{L})(L-\overline{L})^T B^T\Big)=B\tilde{E}\Big((L-\overline{L})(L-\overline{L})^T\Big)B^T\end{aligned}$$

$$\sum\nolimits_X=B\sum\nolimits_L B^T \qquad (6\text{-}15)$$

ملاحظه می شود که برای حالت کلی مدل نیز همان قانون واریانس بدست آید . در حقیقت مدل خطی یک حالت خاصی از حالت عمومی (غیر خطی) می باشد . باید در نظر داشت که ابعاد فیزیکی عناصر تشکیل دهنده ماتریس های B و $\sum_L$ باید بطور صحیح و ضریبی انتخاب شوند که ابعاد فیزیکی صحیح و دلخواه برای عناصر ماتریس $\sum_X$ ایجاد شود .

* بطور صریح اگر داشته باشیم .

$$X=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}x_1(\ell_1,\ell_2,\ldots,\ell_s)\\x_2(\ell_1,\ell_2,\ldots,\ell_s)\\\vdots\\x_n(\ell_1,\ell_2,\ldots,\ell_s)\end{bmatrix}$$

ماتریس B (ماتریس ضرایب) مدل فوق به شکل زیر خواهد بود .

$$\underset{nxs}{B} = \begin{bmatrix} \frac{\partial x_1}{\partial \ell_1} & \frac{\partial x_1}{\partial \ell_2} & \cdots & \frac{\partial x_1}{\partial \ell_s} \\ \frac{\partial x_2}{\partial \ell_1} & \frac{\partial x_2}{\partial \ell_2} & \cdots & \frac{\partial x_2}{\partial \ell_s} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial x_n}{\partial \ell_1} & \frac{\partial x_n}{\partial \ell_2} & \cdots & \frac{\partial x_n}{\partial \ell_s} \end{bmatrix}$$

مثال 6-3

مثال 6-1 را در نظر می گیریم . می خواهیم ماتریس $\sum_X$ را برای قطر d و مساحت α میز مستطیل شکل بدست آوریم . داریم .

$$\sum_L = \begin{bmatrix} S_a^2 & S_{ab} \\ S_{ab} & S_b^2 \end{bmatrix}$$

مدل ریاضی غیر خطی ولی صریح در X است .

$$X = F(L) \qquad \text{یا} \qquad (d,\alpha) = F(a,b)$$

نخست مدل را خطی می کنیم

$$X = (d,\alpha) = (d^\circ,\alpha^\circ) + B[(a,b) - (a^\circ,b^\circ)]$$

جائیکه $(d^\circ,\alpha^\circ) = F(a^\circ,b^\circ)$ و

$$B = \begin{bmatrix} \frac{\partial d}{\partial a} & \frac{\partial d}{\partial b} \\ \frac{\partial \alpha}{\partial a} & \frac{\partial \alpha}{\partial b} \end{bmatrix}$$

$$\frac{\partial d}{\partial a} = \frac{1}{2}\frac{1}{\sqrt{a^2+b^2}}(2a) = \frac{a}{b}, \frac{\partial d}{\partial b} = \frac{b}{d}, \frac{\partial \alpha}{\partial a} = b, \frac{\partial \alpha}{\partial b} = a$$

با جایگزینی مقادیر فوق در ماتریس B داریم :

$$B_{2\times 2} = \begin{bmatrix} \frac{a}{b} & \frac{b}{a} \\ b & a \end{bmatrix}$$

با استفاده از قانون کوواریانس خواهیم داشت .

$$\Sigma_X = \begin{bmatrix} S_d^2 & S_{d\alpha} \\ S_{\alpha d} & S_\alpha^2 \end{bmatrix} = B\Sigma L\ B^T = \begin{bmatrix} \frac{a}{d} & \frac{b}{d} \\ b & a \end{bmatrix} \begin{bmatrix} S_a^2 & S_{ab} \\ S_{\ell a} & S_0^2 \end{bmatrix} \begin{bmatrix} \frac{a}{d} & b \\ \frac{b}{d} & a \end{bmatrix}$$

$$\Sigma_X = \begin{bmatrix} \frac{1}{d^2}(a^2 S_a^2 + 2abS_{ab} + & \frac{1}{d}(ab(S_a^2 + S_b^2) + (a^2+b^2)S_{ab} \\ \frac{1}{d}(ab(S_a^2+S_b^2)+(a^2+b^2)S_{ab} & b^2S_a^2 + 2abS_{ab} + a^2S_b^2 \end{bmatrix}$$

مثال 4-6

فرض می کنیم که چند نمونه ای اولیه (L = (a , b مورد نظر در مثالهای 1-6 و 3-6 داده شده است .

L={a , b }={(128/1,128/1 , 128/2 , 128/0, 128/1) , (62/5 , 62/7 , 62/6 , 62/6 , 62/5)} به سانتیمتر

برآورد آماری از مجهولات مسئله ، (d,α) ، بصورت زیر است .

$$\hat{X} = \begin{bmatrix} \hat{d} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \sqrt{(\bar{a})^2 + (\bar{b})^2} \\ \bar{a}\bar{b} \end{bmatrix}$$

جائیکه $\bar{b},\bar{a}$ عبارتند از برآوردهای آماری (میانگینی) طول و عرض میز مزبور مقادیر $\bar{b},\bar{a}$ از نمونه ها حساب می شود .

$\bar{a} = 128/1$ سانتیمتر و $\bar{b} = 62/59$ سانتیمتر

بنابراین

$$\hat{X} = \begin{bmatrix} \hat{d} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \sqrt{(128/1)^2 + (62/58)^2} \\ 128/1 \times 62/58 \end{bmatrix} = \begin{bmatrix} 124/57 \\ 8016/5 \end{bmatrix}$$

با محاسبه واریانس های b,a ماتریس $\sum_L$ بصورت زیر در می آید

$$\sum_L = \begin{bmatrix} 0/004 & 0 \\ 0 & 0/0056 \end{bmatrix}$$ سانتیمتر مربع

ماتریس فوق نشان می دهد که مؤلفه های b,a را مستقل از همدیگر در نظر گرفته اند . عناصر ماتریس B که در مثال 3-6 داده شده بود بصورت زیر محاسبه می شود .

$$B = \begin{bmatrix} \frac{\bar{a}}{\hat{d}} & \frac{\bar{b}}{\hat{d}} \\ \bar{b} & \bar{a} \end{bmatrix} = \begin{bmatrix} 0/898 & 0/439 \\ 62/58 & 128/1 \end{bmatrix}$$

عناصر ردیف اول بدون بعد فیزیکی و عناصر ردیف دوم دارای بعد سانتیمتر می باشند . در خاتمه ماتریس $\sum_X$ از رابطه زیر محاسبه می شود .

$$\sum x = B\sum l \ B^T = \begin{bmatrix} 0/898 & 0/439 \\ 62/58 & 128/1 \end{bmatrix}\begin{bmatrix} 0/004 & 0 \\ 0 & 0/0056 \end{bmatrix}\begin{bmatrix} 0/898 & 62/58 \\ 0/439 & 128/1 \end{bmatrix}$$

$$= \begin{bmatrix} 0/0043 & 0/5397 \\ 0/5397 & 107/5627 \end{bmatrix}$$ با ابعاد

و از آنجا

$S_d = \sqrt{0/0043} = 0/066$ سانتیمتر

$S_\alpha = \sqrt{107/5627} = 10/37$ سانتیمتر مربع

6-3-2) انتشار خطاها ، حالت نا همبسته (مستقل)

اگر بردار X دارای فقط یک مؤلفه مثلاً x باشد ، ماتریس B در فرمولهای 6-10 یا 6-15 بصورت یک ماتریس با یک سطر و s ستون در خواهد آمد یعنی $B=[B_1, B_2, \ldots, B_S]$ و معادله

$$\sum_x = B\sum_L B^T$$

بصورت فرم کوادراتیک ، با یک سطر و یک ستون یعنی یک عدد ، در می آید .

$$\sum_X = S_x^2$$

اگر علاوه بر مورد فوق ، چند نمونه ای L ناهمبسته فرض شود یعنی $\sum_L$ یک ماتریس قطری بصورت $\sum_L = (S_{j_1}^2, S_{j_2}^2, \ldots, S_{j_s}^2)$ باشد قانون کوواریانس بشکل زیر در می آید

$$\boxed{S_x^2 = \sum^{s} B_i^2 S_i^2} \qquad (6\text{-}18)$$

که در آن $B_x = \dfrac{\partial i}{\partial \ell_i} = \dfrac{\partial x}{\partial \ell_i}$

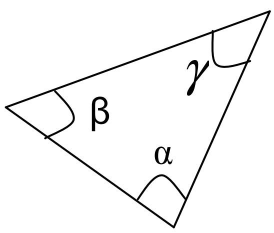
فرمول فوق معروف به قانون انتشار MSE یا بطور ساده قانون انتشار خطا می باشد . قانون پخش یا انتشار خطاها چیزی نیست جز یک حالت خاص از قانون انتشار ماتریس واریانس – کوواریانس قانون پخش خطاها موارد استعمال زیادی در عملیات نقشه برداری و سایر علوم تجربی دارد .

مثال 6-5

در شکل زیر یک مثلث مسطحاتی با زوایای مشاهده شده β, α با مقادیر برآورد شده

$\bar{\alpha} = 32°15'20''$ $\quad S_\alpha = 4''$

$\bar{\beta} = 75°53'32''$ $\quad S_\beta = 3''$



(شکل 6-2)

را در نظر می گیریم و فرض می کنیم که β,α مستقل از همدیگر هستند یعنی $S_{\alpha,\beta}=0$ مطلوبست تعیین زاویه γ همراه با S_γ .

$$\hat{\gamma}=180°-(\overline{\alpha}+\overline{\beta})=72°1'8''$$

$$S_\gamma^2=(\frac{\partial\gamma}{\partial\alpha})^2 S_\alpha^2+(\frac{\partial\gamma}{\partial\beta})^2 S_\beta^2=(-1)(4)^2+(-1)(3)^2=25$$

$$S_\gamma=\partial''$$

مثال 6-6

شکل زیر یک خط ترازیابی را بین دو بنچ مارک C , A همراه با اختلاف ارتفاع های مشاهده شده نشان می دهد . اختلاف ارتفاع هر قسمت که دارای طول ℓ_i می باشد با h_i نشان داده شده است . فرض می کنیم که h_i ها مستقل از همدیگرند و خطای مربعی میانگین (MSE) هر کدام متناسب با طولشان می باشد یعنی $S_{l_i}^2=k\ell_i$ جائیکه K یک عدد ثابت است .

مطلوبست خطای MSE مربوط به اختلاف ارتفاع دو بنچ مارک C , A (ΔH) . برای این کار می نویسیم .

$$\Delta H=H_C-H_A=\sum_{i=1}^{s}h_i$$

مدل ریاضی در اینجا همان معادله فوق یعنی

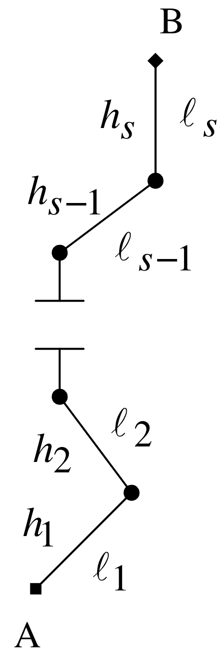
$$\Delta H=h_1+h_2+\cdots+h_s$$

می باشد . از فرمول پخش خطاها ، ΔH MSE را محاسبه می کنیم .

$$S_{\Delta H}^2=(\frac{\partial\Delta H}{\partial h_1})^2 S_{h_1}^2+(\frac{\partial\Delta H}{\partial h_2})^2 S_{h_2}^2+\cdots$$

$$=(1)^2(k\ell_1)+(1)^2(k\ell_2)+\cdots=\sum_{i=1}^{S}k\ell_i=k\sum_{i=1}^{S}\ell_i$$

رابطه فوق نشان می دهد خطای مربعی میانگین ΔH برابر ضریب ثابت k ضربدر طول کلی خط ترازیابی .



شکل 6-3

مثال 3-6 را در نظر بگیرید و فرض کنید که خطاهای مجود در b , a مستقل از همدیگرند یعنی $S_{ab}=0$ در این حالت می توانیم خطاهای d,α را بطور جداگانه محاسبه کنیم (اگر مقصود محاسبه فقط MSE هر کدام باشد) . برای اینکار از قانون انتشار خطاها استفاده می کنیم .

$$S_d^2=(\frac{\partial d}{\partial a})^2 S_a^2+(\frac{\partial d}{\partial \alpha})^2 S_b^2=\frac{1}{d^2}(a^2 S_a^2+b^2 S_b^2)$$

$$S_\alpha^2=(\frac{\partial \alpha}{\partial a})^2 S_a^2+(\frac{\partial \alpha}{\partial \alpha})^2 S_b^2=b^2 S_a^2+a^2 S_b^2)$$

توجه داشته باشید که نتایج فوق همان نتایج بدست آمده در مثال 3-6 می باشد بشرطی که در آنها $S_{ab}=0$ قرار داده شود .

از طرفی اگر بخواهیم کوواریانس $S_{d\alpha}$ را هم حساب کنیم باید از قانون پخش ماتریس واریانس کوواریانس (معادله 15-6) استفاده کنیم که نتیجه زیر بدست خواهد آمد .

$$S_{d\alpha}=\frac{ab}{d}(S_a^2+S_b^2)$$

رابطه فوق نشان می دهد که $S_{d\alpha}\neq 0$ است یعنی کوواریانس بین d,α مخالف صفر است بعبارت دیگر ماتریس $\sum_X$ یک ماتریس ؟ (غیر قطری) بوده هر چند که ماتریس $\sum_L$ قطری بوده است . یعنی عدم همبستگی چند نمونه ای اولیه (a,b) دلیل بر عدم همبستگی چند نمونه ای بدست آمده (d,α) نیست و آنها در هر حالت همبسته هستند .

مطلب فوق نشان دهنده این است که اگر چند نمونه ای بدست آمده (d,α) را بعنوان چند نمونه ای اولیه برای کار بعدی استفاده می کنیم باید ماتریس بدست آمده $\sum_X$ را بصورت کامل آن بعنوان ماتریس $\sum_L$ بعدی استفاده می کنیم . یعنی حق نداریم (d,α) را مستقل فرض می کنیم .

مثال 7-6

مثال 2-6 را دوباره حل کرده اما این بار فرض می کنیم که چند نمونه ای اولیه $L\equiv(\ell_1,\ell_2,\ell_3)$ (d,α)غیر همبسته بوده و دارای ماتریس واریانس ـ کوواریانس

$$\Sigma L=\begin{bmatrix}3&0&0\\0&3&0\\0&0&4\end{bmatrix}$$

باشد . از مثال 2-6 داریم

$$B=\begin{bmatrix}1 & 0 & -3\\ 2 & 1 & 0\end{bmatrix} \quad و \quad X=\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

برای پیدا کردن $\sum_X$ از قانون کوواریانس استفاده می کنیم .

$$\sum_X=\begin{bmatrix}1 & 0 & -3\\ 2 & 1 & 0\end{bmatrix}\begin{bmatrix}3 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 4\end{bmatrix}\begin{bmatrix}1 & 2\\ 0 & 1\\ -3 & 0\end{bmatrix}=\begin{bmatrix}39 & 6\\ 6 & 15\end{bmatrix}=\begin{bmatrix}S_{x_1}^2 & S_x\\ S_{x_2}^2 & S_1^2\end{bmatrix}$$

ماتریس $\sum_X$ دوباره این حقیقت را نشان می دهد که قطری بودن ماتریس $\sum_L$ دلیل بر قطری بودن $\sum_X$ نیست .

از آنجائیکه L غیر همبسته می باشد . واریانسهای x_2,x_1 را می توان بطور جداگانه نیز با استفاده از قانون پخش خطاها ، محاسبه کرد .

$$S_{X_1}^2=(\frac{\partial X_1}{\partial \ell_1})^2 S_{\ell_1}^2+(\frac{\partial X_1}{\partial \ell_2})^2 S_{\ell_2}^2+(\frac{\partial X_1}{\partial \ell_3})^2 S_{\ell_3}^2=(1)^2(3)+(0)^2(3)+(-3)^2(4)=39$$

مقدار فوق همان است که از طریق قانون کوواریانس برای x_1 بدست آمد

مثال 8-6

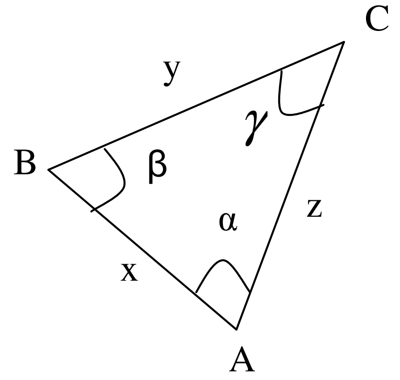
برای تعیین دو ضلع AC = Z و BC = y مثلث مسطحاتی شکل 4-6 طول AB = X و دو زاویه β,α اندازه گیری شدند و مقادیر برآورد شده عبارتند از :

$\bar{X}=10$ متر و $S_X=3$ سانتیمتر

$\bar{\alpha}=90°$ درجه و $S_\alpha=2''$ ثانیه

$\bar{\beta}=45°$ درجه و $S_\beta=4''$ ثانیه

و $S_{X\alpha}=S_{X\beta}$ و ثانیه کمانی^2 $S_{\alpha\beta}=-1$



(شکل 4-6)

مطلوبست برآورد آماری y,z همراه با ماتریس واریانس ـ کوواریانس آنها به سانتیمتر مربع

$$X=(y,z)$$

نخست مدل ریاضی را که مشاهدات را مجهولات مربوط می کند می نویسیم .

$$X=F(L)$$
$$L\equiv(\ell_1,\ell_2,\ell_3)=(\alpha,\beta,\chi),X=(x_1,x_2)=(y,z)$$

از قانون سینوسها در مثلث می توان نوشت

$$\frac{y}{\sin\alpha}=\frac{z}{\sin\beta}=\frac{x}{\sin\gamma}$$

زاویه γجزء مشاهدات نیست یعنی جزو چند نمونه ای اولیه نیست بنابراین آن را بر حسب زوایای مشاهده شده قرار میدهیم یعنی $\sin\gamma=\sin(k+A)$ پس خواهیم داشت .

$$y=\frac{x\sin\alpha}{\sin(\alpha+\beta)}$$

$$z=\frac{x\sin\beta}{\sin(\alpha+\beta)}$$

در فرمولهای فوق مقادیر برآورد شده مشاهدات را قرار داده مجهولات را برآورد می کنیم .

$$\hat{y}=10\sqrt{2}\approx 14/14 \text{ متر}$$

$$\hat{z}=10 \text{ متر}$$

مدل ریاضی شکل زیر خواهد بود

$$x=\begin{bmatrix} y \\ z \end{bmatrix}=\begin{bmatrix} y(\alpha,\beta,x) \\ z(\alpha,\beta,\chi) \end{bmatrix}=\begin{bmatrix} x\sin\alpha/\sin(\alpha+\beta) \\ x\sin\beta/\sin(\alpha+\beta) \end{bmatrix}$$

برای محاسبهٔ $\sum_x=B\sum_L B^T$ نخست ماتریس B را معین می کنیم .

$$B=\begin{bmatrix} \frac{\partial y}{\partial\alpha} & \frac{\partial y}{\partial\beta} & \frac{\partial y}{\partial\chi} \\ \frac{\partial z}{\partial\alpha} & \frac{\partial z}{\partial\beta} & \frac{\partial z}{\partial\chi} \end{bmatrix}=\begin{bmatrix} z/\sin(\alpha+\beta) & -y/tag(\alpha+\beta) & \frac{y}{x} \\ -z/tag(\alpha+\beta) & y/\sin(\alpha+\beta) & \frac{z}{x} \end{bmatrix}$$

و ماتریس $\sum_L$ از مقادیر داده شده برای واریانس ها بصورت زیر تنظیم می شود .

$$\Sigma L=\begin{bmatrix} S_{\alpha}^{2} & S\chi\beta & S_{\alpha\chi} \\ S_{\beta\alpha} & S_{\beta}^{2} & S_{\beta\chi} \\ S\chi\alpha & S_{\chi\beta} & S_{\chi}^{2} \end{bmatrix}=\begin{bmatrix} 4 & -1 & 0 \\ -1 & 16 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

خیلی مهم است که ترتیب قرار گرفتن عناصر در دو ماتریس را نظیر ترتیب عناصر در چند نمونه ای اولیه انتخاب کرد در غیر اینصورت عناصر ماتریس $\sum_L$ و B بدون مفهوم خواهند بود .

حال ابعاد فیزیکی عناصر دو ماتریس $\sum_L$ و B را موافق همدیگر و طوری انتخاب می کنیم که نتیجه ، یعنی $\sum_X$ به سانتیمتر مربع باشد برای این کار ماتریس B را به شکل زیر در می آوریم .

$$B=\begin{bmatrix} z/P''\cos(\alpha+\beta) & -y/p''\sin(\alpha+\beta) & \frac{y}{x} \\ -z/p''tag(\alpha+\beta) & y/p''\sin(\alpha+\beta) & \frac{z}{x} \end{bmatrix}$$

جاییکه ثانیه کمانی $p'' = 206265 = 2\times10^5$ می باشد .

با جایگزینی مقادیر برآورد شده کمیتهای d, z, μ, α, x ماتریس فوق خواهیم داشت

$$B = \begin{bmatrix} 0/007 & 0/007 & 1/414 \\ 0/005 & 0/010 & 1/000 \end{bmatrix}$$

و ماتریس $\sum_x$ بصورت زیر محاسبه می شود

$$\Sigma x = \begin{bmatrix} 0/007 & 0/007 & 1/414 \\ 0/005 & 0/010 & 1/000 \end{bmatrix} \begin{bmatrix} 4 & -1 & 0 \\ -1 & 16 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} 0/007 & 0/005 \\ 0/007 & 0/010 \\ 1/414 & 1/000 \end{bmatrix}$$

$$= \begin{bmatrix} 18/0009 & 12/7272 \\ 12/7272 & 9/0016 \end{bmatrix} = \begin{bmatrix} 18 & 13 \\ 13 & 9 \end{bmatrix}$$ سانتیمتر مربع

و از آنجا

$S_y = \sqrt{18} = 4/2$ سانتیمتر

$S_x = \sqrt{9} = 3$ سانتیمتر

با توجه به نتایج فوق و با در نظر گرفتن ماتریسB معلوم می شود که دقت زیاد در اندازه گیری زوایای β, α تاثیر چندانی در بهبودی برآوردهای y , z نخواهد داشت اما دقت اندازه گیری ضلع x تاثیر زیادی در دقت مقادیر بدست آمده برای y , z خواهد داشت .

بنابراین از قانون پخش خطاها می توان در تشخیص عامل یا عوامل موثر (مشاهده یا مشاهدات) در برآورد مجهولات استفاده کرد بعد از مشخص کردن این مشاهده یا مشاهدات داخل سری مشاهدات (چند نمونه ای اولیه) می توان دقت اندازه گیری آن را چنان بالا برد که نتایج بدست آمده برای مجهولات دارای دقت دلخواه باشند این عمل را « آنالیز اولیه» می نامند .

آنالیز اولیه قبل از اینکه اندازه گیری واقعی شروع شود انجام می گیرد . برای آنالیز اولیه مقـادیر تقریبی از مجهولات کافی است .

عمل آنالیز اولیه منحصر به تعیین مشخصات تکنیکهای مختلف اندازه گیری می شود که طی آن دقتهای دلخواه برای مجهولات حاصل می شود جزئیات بیشتر در مورد آنالیز اولیه در بخش 6-3-5 خواهد آمد .

6-3-3 انتشار خطاهای غیر اتفاقی ، انتشار خطای کلی

نیاز به پیشگویی یا تخمین اولیه قدر مطلقMSE (بسیاری از خطاهای اتفاقی) یک تابع از مشاهدات ـ این اساساً کاری از که قانون بخش خطاها انجام می دهد – در مورد خطاهای غیر اتفاقی نیز احساس می شود . خطاهای غیر اتفاقی را گاهی خطاهای سیستماتیک نیز می گویند و

قانون حاکم بر رفتار آنها مشخص نیست . بنابراین برای آنالیز اولیه تعیین مشخصات اندازه گیری مقادیری برای این خطاهای غیر اتفاقی که در اندازه گیریهای پی در پی ظاهر می شوند در نظر گرفته می شوند .

مسئله ممکن است به شکل زیر مطرح شود . فرض کنیم که مدل ریاضی

$$x=f(L) \qquad (6\text{-}19)$$

را داریم که در آنن x یک کمیت تنها و f یک تابع از چند نمونه ای L که یک بردار $L=(\ell_1,\ell_2,\dots,\ell_s)$ از کمیتهای مختلف مشاهده شده است می باشد .

فرض می کنیم که L غیر هم بسته نیــز باشـد. مـا مـی خـواهیم تـأثیر خطاهـای کوچـک غیـر اتفاقی $\delta\ell_i$ هر مشاهده ℓ_i را روی x تعیین کنیم این اثر را روی x با δ_x نشان می دهیم .

مسئله را می توان آناً با استفاده مجدد از بسط سری تیلور حل نمود تابع x را حـول یـک نقطـه تقریبی به سری تیلور بسط داده و دو جملهٔ اول آن را فقط در نظر می گیریم .

$$x=f(L^\circ)+\left.\frac{\partial f}{\partial L}\right|_{L=L_0}(L-L^\circ)$$

$$=x_L^0+\sum_{i=1}^{s}\left.\frac{\partial f}{\partial \ell}\right|_{\ell=\ell^\circ}(\ell-\ell^\circ) \qquad (6\text{-}20)$$

با جایگزینی $\delta\ell_i$ به جای $(\ell_i-\ell_i^0)$ و به جای $(x-x^\circ)$ ، δx در فرمول 20-6 خواهیم داشت .

$$\boxed{\delta_x=\sum_{i=1}^{s}\left.\frac{\partial f}{\partial \ell_i}\right|_{\ell_i=\ell_i^2}\delta\ell_i} \qquad * \quad (6\text{-}21)$$

که فرمول انتشار خطاهای غیر اتفاقی است .

توجه داشته باشیم که علامت های هر دو مشتق جزئی $\delta\ell_i,\frac{\partial f}{\partial \ell_i}$ باید در نظر گرفته شود (این فرمول را با فرمول 18-6 مقایسه کنید) .

ممکن است سوال شود چه خطائی در x ایجاد می شود اگر مـشاهدات ℓ_i دارای هـر دو خطـای اتفاقی و غیر اتفاقی باشند در عین حالتی ما خطای کل را به صورت زیر تعریف می کنیم .

$$T=\sqrt{\delta^2+S^2} \qquad (6\text{-}22)$$

که در آن δ خطای غیر اتفاقی و S خطای اتفاقی(RMS) می باشد .

تأثیر هر دو خطا در x که با فرمول فوق داده شده و با استفاده از فرمولهای (18-6) و (21-6) به صورت زیر در می آید .

$$T_x=\sqrt{(\sum_{i=1}^{s}\frac{\partial f}{\partial \ell_i}\partial \ell_i)^2+\sum_{i=1}^{s}(\frac{\partial f}{\partial \ell_i})^2 S_{\ell_i}^2}$$
$$=\sqrt{\sum_{i=1}^{s}(\frac{\partial f}{\partial \ell_i})^2(\partial \ell_i^2+S_{\ell_i}^2)\ +\sum_{i=1}^{s}\frac{\partial f}{\partial \ell_i}\frac{\partial f}{\partial \ell_i}\delta\ell_i\delta\ell_j}$$
$$T_x=\sqrt{\sum_{i=1}^{s}(\frac{\partial f}{\partial \ell_i})^2 T_i^2+q} \qquad (6\text{-}23)$$

جاییکه q مثل 1- کوواریانس بین خطاهای غیر اتفاقی و T_i خطای کل هر کدام از مشاهدات می باشد .

* برای حفظ ارزش تابع در بسط آن به سری تیلور باید در نظر داشت که کوچک بودن خطای $\delta\ell_i$ نسبت به ℓ_i ضروری است .

به طوری که در بخش 2-4 گفته شد خطاهای غیر اتفاقی (سیستماتیک) ممکن است شناخته شده و یا تابع یک سری پارامتر باشند در این صورت تاثیر آنها در δ_x, x می توانند بر حسب پارامترهای مذکور فرموله شود .

مثال 6-9

مثال 6-2 را دوباره در نظر می گیریم و فرض می کنیم که چند نمونه ای اولیه غیر هم بسته بوده و دارای ماتریس واریانس و کوواریانس

$$\sum L=\begin{bmatrix}3&0&0\\0&3&0\\0&0&4\end{bmatrix}$$

و همچنین دارای خطای سیستماتیک

$$\delta L\equiv(\delta\ell_1,\delta\ell_3,\delta\ell_3)=(-1/5,2,0/5)$$

باشند واحد خطاهای فوق نظیر واحد خطاهای اتفاقی داده شده است .

مطلوبست تعیین خطاهای کل در برآورد کمیتهای x_2, x_1 مدل ریاضی همان است در مثال 6-2 داده شده است .

خطای کل به وسیلهٔ فرمول 6-22 محاسبه می شود .

$$T_{x_1}=\sqrt{(\delta_{x_1}^2+S_{x_1}^2)}$$
$$T_{x_2}=\sqrt{(\delta_{x_2}^2+S_{x2}^2)}$$

و داریم

$$S_{x_1}^2 = \sum_{i=1}^{3} (\frac{\partial x_1}{\partial \ell_i})^2 S_{\ell_i}^2 = 39, S_{x_2}^2 = \sum (\frac{\partial x_2}{\partial \ell_i})^2 S_{\ell_i}^2 = 15$$

و خطاهای δ_{x_1} و δ_{x_1} ناشی از خطاهای غیر اتفاقی ، از فرمول 6-21 محاسبه می شوند .

$$\delta_{x_1} = \sum_{i=1}^{3} \frac{\partial x_1}{\partial \ell_i} S\ell_i = (1)(1/5) + (0)(2) + (-3)(0/5) = -3$$

$$\delta_{x_1} = \sum_{i=1}^{3} \frac{\partial x_2}{\partial \ell_i} S\ell_i = (2)(1/5) + (1)(2) + (0)(0/5) = -1$$

$$T_{x_1} = \sqrt{(-3)^2 +} = 6/93 \quad \text{و} \quad T_{x_2} = \sqrt{(-1)^2 + 15} = 4$$

مثال ۱۰-6

مثال 6-6 را در نظر می گیریم. علاوه بر اطلاعات داده شده در آنجا فرض می کنیم که اختلاف ارتفاع h_i دارای خطای سیستماتیک و بصورت $\delta_{h_i} = k'h_i$ می باشد. k' یک ضریب ثابت می باشد. مطلوبست تعیین خطای کل روی ΔH .

$$\Delta H = H_C - H_A = \sum_{i=1}^{s} h_i = h_1 + h_2 + ... + h_s$$

خطای کل از فرمول زیر بدست می آید:

$$T_{\Delta H} = \sqrt{\delta_{\Delta H}^2 + S_{\Delta H}^2}$$

در مثال 66-6 مقدار $S_{\Delta H}^2$ بدست آمد :

$$S_{\Delta H}^2 = k\ell_{AC}$$

جائیکه k یک ضریب ثابت و $\ell_{AC} = \sum_{i=1}^{s} \ell_i$ بود. $\delta_{\Delta H}$ بصورت زیر محاسبه می شود :

$$\delta_{\Delta H} = \sum_{i=1}^{s} \frac{\partial \Delta H}{\partial h_i}$$

در حالیکه داریم :

$$\frac{\partial \Delta H}{\partial h_1} = \frac{\partial \Delta H}{\partial h_2} = ... = \frac{\partial \Delta H}{\partial hs} = 1 \quad \text{و} \quad \delta_{h_i} = k'h_i$$

پس

$$\delta_{\Delta H} = \sum_{i=1}^{s} k'h_i = k'\sum_{i=1}^{s} h_i = k'\Delta H$$

و از آنجا خطای کل روی ΔH بصورت زیر خواهد بود :

$$T_{\Delta H} = \sqrt{k'^2 \Delta H^2 + k\ell_{AC}}$$

٤-٣-6 حذف و گرد کردن اعداد

در هر محاسبه مجبوریم اعدادی را که با آنها سر و کار داریم بصورتی نشان دهیم . اعداد حقیقی ممکن است غیر قابل نمایش با کسر متعارفی ، نظیر اعداد π و e و $\sqrt{2}$ و یا اعدادی $\frac{1}{3}$ یا ارقام اعشاری نامحدود می باشند. در نمایش چنین اعدادی به تعداد محدود و مورد نیاز از ارقام اعشاری قناعت کرده و بقیه را حذف می کنیم.

حذف ارقام اعشاری اساساً به دو طریق صورت می گیرد. یا ارقام اعشاری مازاد بر احتیاج را حذف کرده و باقیمانده را بدون هیچ تغییری بجای عدد اصلی قبول می کنیم یا بعد از حذف کردن ، اولین رقم اعشاری حذف نشده را با توجه به ارقام حذف شده گرد می کنیم (یعنی صفر یا یک را به آن اضافه می کنیم).

اولین طریق را می توان بصورت زیر نشان داد :

$$a = a_T = \frac{Int(a \cdot 10^n)}{10^n} \qquad (6-٢٤)$$

جائیکه a عدد اصلی بوده و n تعداد ارقام اعشاری مورد نیاز بعد از ممیز و Int مخفف کلمه Integer یعنی قسمت صحیح (صرفنظر از ارقام اعشاری) عدد می باشد.

مثال ١١-6

عدد $\pi = 3/141592...$ را تا ٣ رقم اعشار بنویسید $(\pi = 3)$.

$$\pi = \pi_T = Int(\pi \cdot 10^3)/10^3 = Int(3141/592)/10^3 = \frac{3141}{10^3} = 3/141$$

دومین طریق را می توان بصورت زیر نشان داد :

$$a = a_R = \frac{Int(a \cdot 10^n + 0.5)}{10^n} \qquad (6-٢٥)$$

مثال ١٢-6

عدد π را تا سه رقم $(\pi = 3)$اعشار گرد کنید.

$$\pi = \pi_R = Int(\pi \cdot 10^3 + 0.5)/10^3 = Int(3141/59... + 0.5)/10^3$$
$$= Int(3142/09...)/10^3 = \frac{3142}{10^3} = 3/142$$

ملاحظه کنید که خطای ناشی از حذف ارقام اعشاری در دو طریق فوق مختلف می باشد. اگرخطای ناشی از حذف (Truncation) ارقام اعشاری را در"a" با δ_{a_T} و خطای ناشی از گرد کردن (Rounding) عدد "a" را با δ_{a_R} نشان دهیم خواهیم داشت :

$$\delta_{a_T} = a - a_T \in [0, 10^{-n}]$$

$$\delta_{a_R} = a - a_R \in [-0.5\times10^{-n}, 0.5\times10^{-n}]$$

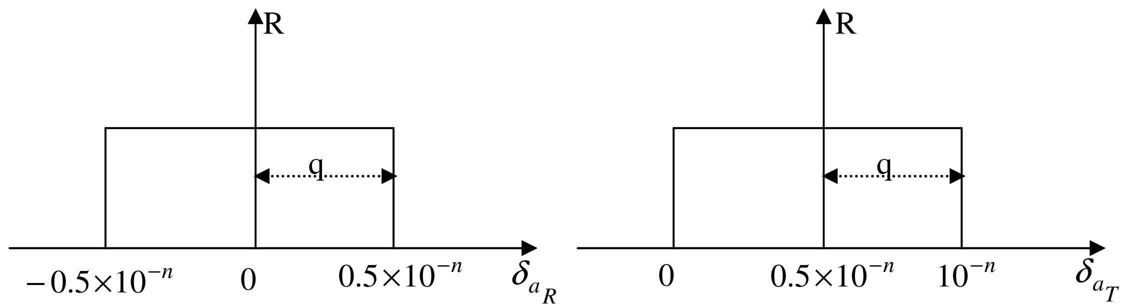
ممکن است فرض کنیم که δ_{a_T} دارای یک متغیر تصادفی که برحسب تابع PDF مستطیل شکل (یکنواخت) پخش شده است ، بخش ۵-2-۳ .

$$R(0.5\times10^{-n}, \sigma, \delta_{a_T}) \qquad (6\text{-}۲6)$$

و δ_{a_R} دارای تابع پخش مادر

$$R(0, \sigma, \delta_{a_R}) \qquad (6\text{-}۲۷)$$

می باشد. دو تابع فوق در شکل زیر مشخص شده اند.



خطاي گرد كردنPDF تابع — **خطاي حذف كردنPDF تابع**

شکل ۵-6

از مثال ۱۷-۳ در بخش ۵-۲-۳ می دانیم $\sigma = q/\sqrt{3}$ به طوریکه q نصف طول مستطیل بود. در این حالت $q = 0.5\times10^{-n}$ می باشد. در نتیجه خواهیم داشت $\sigma \cong 0/28\times10^{-n}$

بعلت مختلف بودن میانگین های دو خطا ، خطای حذف کردن ، برحسب " قانون پخش خطای کل " و خطای گرد کردن ، برحسب " قانون پخش خطای اتفاقی " انتشار می یابد. بنابراین اگر عدد x تابع چند عدد دیگر بشکل زیر باشد :

$$x = f(L) \qquad (6\text{-}۲۸)$$

$$L = (\ell_i), i = 1,2,,...,S$$

L عبارتست از یک سری اعداد که یا بطریق حذفی بدست آمده اند و یا گرد شده اند. میتوانیم خطای ناشی از عمل حذف کردن و گرد کردن اعداد را روی x پیدا کنیم.

$$\delta_{x_T} = \sqrt{(\sum_{i=1}^{s}\frac{\partial f}{\partial \ell_i}0.5\times 10^{-n})^2 + \sum_{i=1}^{s}(\frac{\partial f}{\partial \ell_i})^2\frac{10^{-2n}}{12}} \qquad (6\text{-}۲۹)$$

$$\delta_{x_R} = \sqrt{\sum_{i=1}^{s}(\frac{\partial f}{\partial \ell_i})^2\frac{10^{-2n}}{12}} \qquad (6\text{-}۳۰)$$

دو فرمول فوق نشان می دهند که خطای ناشی از گرد کردن اعداد ℓ_i روی x بمراتب کمتر از خطای ناشی از عمل حذف کردن می باشد. بدین جهت است که همیشه اعداد را گرد کرده بکار می بریم.

مثال ۱۳-6

مطلوبست تعیین خطای روی x (مجموع هزار عدد a_i)

$$x = \sum_{i=1}^{1000} a_i$$

الف- اگر هر کدام از اعداد a_i دارای پنج رقم اعشار بوده و بطریق حذفی بدست آمده باشند.

ب- اگر هر کدام از a_i ها تا رقم پنجم اعشاری گرد شده باشند.

راه حل

الف- خطای δ_{x_T} ناشی از حذف ارقام اعشاری بعد از رقم پنجم از فرمول ۲۹-6 محاسبه می شود :

$$\delta_{x_T} = \sqrt{(\sum_{i=1}^{1000}\frac{\partial x}{\partial a_i}0.5\times 10^{-5})^2 + \sum_{i=1}^{1000}(\frac{\partial x}{\partial a_i})^2\frac{10^{-10}}{12}} = \sqrt{0.25\times 10^{-4} + \frac{10^{-5}}{12}}$$

$$= 0.005001 \cong 0..005$$

ب- خطای δ_{x_R} ناشی از گرد کردن a_i ها تا رقم پنجم اعشاری از فرمول ۳۰-6 محاسبه می شود :

$$\delta_{x_R} = \sqrt{\sum_{i=1}^{1000}(\frac{\partial x}{\partial a_i})^2\frac{10^{-10}}{12}} = \sqrt{1000(\frac{10^{-10}}{12})} \cong 0.00091$$

ملاحظه می شود که خطای ناشی از گرد کردن اعداد a_i روی x بمراتب کمتر از خطای ناشی از حذف کردن می باشد.

۵-۳-6 حدود تالرانس ، مشخصات و آنالیزهای اولیه

یکی دیگر از کاربردهای مهم " قانون پخش خطاها " عبارتست از تعیین مشخصات اندازه گیری یک سری از کمیتهای فیزیکی که منجر به برآورد آماری یک سری دیگر از کمیتهای فیزیکی می شوند ، وقتیکه ماکزیمم خطای مجاز (یا حدود تالرانس) در کمیتهای برآورد شده (نتایج) از قبل معلوم می باشند. عمل فوق را آنالیز اولیه می نامند.

تعیین مشخصات ، منجر به طرح صحیح برنامه مشاهدات ، یعنی انتخاب تکنیک مشاهدات و دستگاه اندازه گیری و غیره می شود. بطوریکه نتایج بدست آمده دارای خطای مجاز (حدود قابل قبول تالرانس) خواهند بود.

در تعیین مشخصات ، باید هر دو خطاهای اتفاقی و خطاهای اجتناب ناپذیر غیر اتفاقی (سیستماتیک) در نظر گرفته شوند.

معمولاً خواسته می شود که مشخصات و دستورالعمل ، چنان تعیین گردند که نتایج بدست آمده دارای خطای مجاز با احتمال ۹۹ درصد باشند. اگر چنانچه خطاهای اتفاقی دارای تابع PDF گوس می باشند ، نتایج واقعی نباید دارای خطای کل (متشکل از خطای اتفاقی δ و۲/۵ تا ۳ برابر خطای RMS ، ضریب ۲/۵ تا ۳ مربوط به احتمال ۹۹ درصد می شود) بزرگتر از حد قابل قبول باشند. یعنی

$$\sqrt{\delta^2 + (3\sigma)^2} \le \varepsilon \qquad (6\text{-}۳۱)$$

مثال ۱٤-6

فرض کنید که می خواهیم فاصله $D = 1000$ متر را با یک خطای نسبی بهتر از 10^{-4} و با استفاده از یک نوار بیست متری اندازه گیری کنیم. نوار بیست متری در مقایسه با یک استاندارد ، دقتی کمتر از $3\sigma < 1^{mm}$ ، یعنی حدود تالرانس $\pm 1^{mm}$ نشان داد. فرض می کنیم که طول D به پنجاه دهنه $d_c, c = 1,2,...,50$ حدود بیست متری تقسیم شده است. هر دهنه فقط دو بار (رفت F_i و برگشت B_i) اندازه گیری شده است. برای رسیدن به خطای نسبی فوق (10^{-4}) چه اختلافی در رفت و برگشت هر دهنه قابل قبول است؟

راه حل :

حدود تالرانس در D خطای کل مجاز در D عبارتست از :

$$T_D = 1000 \times 10^{-4} = 0.1mm = 10cm$$

خطای کل T_D مرکب از دو خطای سیستماتیک δ_D و خطای اتفاقی σ_D بصورت زیر میباشد :

$$T_D = \sqrt{\delta_D^2 + (3\sigma_D)^2}$$

ضریب ۳ در فرمول فوق ، برای احتمال ۹۹ درصد می باشد. بنا به تابع PDF گوس مفروض و با دانستن این که

$$D = \sum_{i=1}^{50} d_i \quad \text{و} \quad d_i = \frac{1}{2}(F_i + B_i)$$

می توان نوشت :

$$\delta_D = \sum_{i=1}^{50} \frac{\partial D}{\partial d_i} \delta_{d_i} \quad \text{و} \quad \delta_{d_i} = \frac{\partial d_i}{\partial F_i}\delta_{F_i} + \frac{\partial d_i}{\partial B_i}\delta_{B_i} = \frac{1}{2}\delta_{F_i} + \frac{1}{2}\delta_{B_i}$$

$$\delta_{d_i} = \frac{1}{2}(\delta_i + \delta_i) = \delta_i$$

جائیکه $\delta_i \le 1^{mm}$ (خطای سیستماتیک نوار متر با احتمال ۹۹ درصد) . بنابراین داریم :

$$\delta_D = \sum_{i=1}^{50} \delta_i \le 50^{mm} = 5^{cm}$$

با جایگزینی مقدار δ_D در فرمول خطای کل D خواهیم داشت :

$$(3\sigma_D)^2 \le T_D^2 - \delta_D^2 = (10)^2 - (5)^2 = 75cm^2$$

$$\sigma_D \le \frac{75}{9} \quad \text{و} \quad \sigma_D^2 \le 8.33cm^2$$

که عبارتست از حد مجاز خطای MSE در طول D ، برای رسیدن به دقت خواسته شده. اگر خطای MSE هر دهنه را با $\sigma_{d_i} = \sigma_d$ (برای تمام دهنه ها یکسان در نظر می گیریم) نشان دهیم ، خواهیم داشت :

$$\sigma_D = \sum_{i=1}^{50} \sigma_d^2 = 50\sigma_d^2$$

$$\sigma_d^2 = \frac{\sigma_D^2}{50} \Rightarrow \sigma_d^2 \le \frac{8.33}{50} = 0.16^{cm^2}$$

با در نظر داشتن اینکه هر دو دهنه d_i از فرمول $d_i = \frac{1}{2}(F_i + B_i)$ محاسبه می شود و با فرض اینکه خطای MSE هر رفت و یا برگشت (بطور مساوی) σ^2 می باشد داریم :

$$\sigma_{d_i}^2 = \left(\frac{\partial d_i}{\partial F_i}\right)^2 \sigma_{F_i}^2 + \left(\frac{\partial d_i}{\partial B_i}\right)^2 \sigma_{B_i}^2 = \left(\frac{1}{2}\right)^2 \sigma^2 + \left(\frac{1}{2}\right)^2 \sigma^2 = \frac{1}{2}\sigma^2$$

و از آنجا

$$\sigma^2 = 2\sigma_d^2 \quad \text{و} \quad \sigma^2 \le 2(0.1\overline{6}) = 0.3\overline{3}^{cm^2}$$

در صورت مسئله خواسته شده است که برای رسیدن به دقت 10^{-4} در اندازه گیری D ، چه اختلافی در رفت و برگشت یک دهنه قابل قبول است.
اگر این اختلاف را با $\Delta_i = F_i - B_i$ نشان دهیم ، خواهیم داشت :

$$\sigma_{\Delta_i}^2 = \sigma_{\Delta}^2 \left(\frac{\partial \Delta_i}{\partial F_i}\right)^2 \sigma_{F_i}^2 + \left(\frac{\partial \Delta_i}{\partial B_i}\right)^2 \sigma_{B_i}^2 = (1)^2 \sigma^2 + (-1)^2 \sigma^2 = 2\sigma^2$$

$$\sigma_{\Delta}^2 = 2\sigma^2 \quad \text{و} \quad \sigma_{\Delta}^2 \le 2(0.33) = 0.66^{cm^2}$$

و در نتیجه

$$\sigma_{\Delta} \le 0.8^{cm}$$

اگر فرض کنیم که Δ ها مطابق تابع PDF گوس پخش شده اند ، معنی $\sigma_{\Delta} \le 0.8^{cm}$ اینست که خطای استاندارد اختلافهای رفت و برگشت باید کمتر و یا مساوی با ۰/۸ سانتیمتر باشد. در نتیجه مشخصات و دستورالعمل اندازه گیری طول D بصورت زیر داده می شود. 6۸ درصد اختلافهای رفت و برگشت دهنه ها باید بین ± 0.8 سانتیمتر و ۹۵ درصد اختلافها بین ± 1.6 سانتیمتر باشند. ملاحظه می شود که یک عامل با تجربه ، درعمل ، به راحتی با مشخصات فوق اندازه گیری می کند.

6-٤ مسئله سرشکنی

6-٤-١ فرموله کردن مسئله

در بخش 6-۲ انتقال آماری را تعریف کرده و گفتیم که مسئله سرشکنی عبارتست از انتقال آماری :

$$\left(\overline{L}, \sum L\right) \rightarrow \left(\widehat{X}, \sum X\right) \qquad (6\text{-}۳۲)$$

برای مدل ریاضی با اطلاعات بیش از مورد نیاز

$$F(L, X) = 0 \qquad (6\text{-}۳۳)$$

منظور از مدل با اطلاعات بیش از مورد نیاز مدلی است که در آن بردار $\overline{L}$ دارای تعداد مؤلفه های زیادتر از مورد نیاز مدل باشد که این حالت منجر به تعداد معادلات بیش از مجهولات خواهد شد. چنین مدلی دارای تعداد زیادی جواب برای X خواهد بود. تنها راه حل چنین مدل یا تنها راه ارضای تمام روابط موجود در مدل ، اینست که به بعضی و یا تمام مؤلفه های $\overline{L}$ اجازه داده شود که مختصری تغییر کرده ، بطوریکه یک جواب برای $\widehat{X}$ حاصل شود. بعبارت دیگر

بردار $\overline{L}$ را بعنوان یک مقدار تقریبی از بردار $\widehat{L}$ (که راه حل یکسان $\widehat{X}$ را حاصل می شود) در نظر گرفته ، سپس آخرین مقادیر $\widehat{L}$ و $\widehat{X}$ را برای هر دو بردار L و x جستجو کرد.

با قرار دادن

$$\widehat{L}-\overline{L}=V \qquad (6\text{-}٣٤)$$

مدل ریاضی ٣٣-6 را می توان دوباره بصورت زیر نوشت :

$$F\left(\widehat{L},X\right)=F\left(\overline{L}+V,X\right)=0 \qquad (6\text{-}٣٥)$$

که در آن V را بردار اختلافات می نامند.

ملاحظه کنید که بردار V در اینجا خیلی شباهت به مقادیر v ها در بخش ٨-٤ دارد. از نقطه نظر ریاضی اختلاف زیادی بین V و v ها وجود ندارد. ولی طبق تعاریف آنها ، V برداری است که اختلافها در s تا کمیت جداگانه بوده ولی v ها عبارتند از اختلافها در n مشاهده از یک کمیت فیزیکی برای نشان دادن تعادل ریاضی در بردار فوق . در بخش بعدی ، محاسبات میانگین یک نمونه را بعنوان یک مسئله سرشکنی در نظر می گیریم.

٢-٤-6 میانگین نمونه بعنوان یک مسئله آموزشی سرشکنی ، وزن ها

یک نمونه تصادفی $L=\left(\ell_1,\ell_2,...,\ell_n\right)$ از n مشاهده یک کمیت فیزیکی را در نظر میگیریم. مجموعه تعاریف فوق را بصورت $\overline{L}=\left(\overline{\ell}_1,\overline{\ell}_2,...,\overline{\ell}_n\right)$ که دارای m مشاهده مشخص از ℓاست ، نشان می دهیم. فرض کنید که می خواهیم در مدل ریاضی زیر یک برآورد آماری برای x پیدا کنیم.

$$x=\ell \qquad (6\text{-}٣6)$$

بدیهی است که مدل دارای اطلاعات بیش از مورد نیازاست. برای اینکه کمیت های $\overline{\ell}_j, j=1,2,...,m$ با هم مساوی نبوده و مختلف از همدیگر می باشند و در نتیجه جوابهای مختلف برای مدل فوق را بصورت زیر تغییر می دهیم :

$$x=\overline{\ell}_j+v_j \qquad j=1,2,...,m \qquad (6\text{-}٣٧)$$

جائیکه v_jها همان اختلافات می باشند. مسئله ای که در اینجا مطرح شده است همان است که در بخش ٧-٤ مطرح بود. ولی راه رسیدن به نتیجه ، مختصری فرق دارد. علت انتخاب این راه ، بخاطر مشابهت های آن با مسائل بعدی می باشد. برای محاسبه میانگین ، در بخش ٧-٤ ، تمام n مشاهده از کمیت ℓ را در نظر می گیریم. در حالیکه در اینجا تعداد m مشاهده مشخص از همان کمیت را در نظر می گیریم. m مشاهده فوق ، یعنی $\overline{\ell}_j, j=1,2,...,m$ که مجموعه $\overline{L}$* را تشکیل می دهند.

بنابراین برای محاسبه میانگین $\overline{\ell}$ باید از فرمول دوم که در بخش ۳-۱-۳ معرفی شد استفاده می کنیم :

$$\overline{\ell} = \sum_{j=1}^{m} \overline{\ell}_j P\left(\overline{\ell}_j\right) = \sum_{j=1}^{m} \overline{\ell}_j P_j \qquad (6\text{-}38)$$

* در اینجا نمونه $\overline{L} = \left(\overline{\ell}_1, \overline{\ell}_2, ..., \overline{\ell}_n\right)$ را می توان یک نمونه از گردکردن مشاهدات در نظر گرفت. یعنی هر مؤلفه $\overline{\ell}_j, j = 1,2,...,m$ شامل C_j (فرکانس) مشاهده مربوطه در نمونه اصلی می باشند.

بنا به بخش ۳-۱-۳ داریم $P_j = \frac{C_j}{n}$ که درآن C_j عبارتست از تعداد ℓ_j ها در نمونه اصلی L که دارای n مشاهده می باشد و P_j همان احتمال تجربی یا واقعی می باشد. بعبارت دیگر اگر می خواهیم که $\hat{x}$ برابر $\overline{\ell}$ (میانگین) باشد ، مدل ۳۷-6 جواب زیر را می دهد :

$$\hat{x} = \sum_{j=1}^{m} \overline{\ell}_j P_j \qquad (6\text{-}39)$$

و یا بصورت ماتریسی

$$\hat{x} = P^T \overline{L} \qquad (6\text{-}40)$$

که در آن $P^T = [P_1, P_2, ..., P_m]$ می باشد. ضرایب P_j ها را ضرایب وزن یا بطور ساده وزن ها می نامند و $\hat{x}$ را میانگین وزن دار (اسمی که از مکانیک گرفته شده است) می نامند. ملاحظه کنید که وزن ها چیزی نیستند جز احتمال های تجربی. وزن زیاد به مقادیری داده می شود که بیشتر قابل اعتماد هستند. یعنی بیشتر در نمونه اصلی ظاهر شده اند.

میانگین نمونه یک " برآورنده کمترین مربعات " می باشد. بدین معنی که مجموع مربعات اختلاف ها ، v_i ها ، رابطه زیر، بازای $\ell_0 = \overline{\ell}$ می نیمم می باشد :

$$\frac{1}{n}\sum_{i=1}^{n} V_i^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\ell_i - \ell^0\right)^2 \qquad (6\text{-}41)$$

برای اثبات آن ، از رابطه فوق که تابعی از ℓ^0 می باشد نسبت به ℓ^0 مشتق گرفته و حاصل را مساوی صفر قرار می دهیم (شرط می نیمم) . از حل این معادله مزبور مقدار $\ell^0 = \frac{1}{n}\sum_{i=1}^{n}\ell_i = \overline{\ell}$ بدست می آید. بنابراین می توان گفت که میانگین یک نمونه ، واریانس می نیمم را برای نومنه ایجاد می کند. اگر مجموعه تعاریف نمونه را در نظر بگیریم ، شرط می نیمم و رابطه ٤١-6 بصورت زیر نوشته می شود :

$$\min_{\ell^0 \varepsilon R}\left[\sum_{j=1}^{n} P_j v_j^2\right] \qquad (6\text{-}٤٢)$$

که در آن $v_j = \bar{\ell}_j - \ell^0$ می باشد.

رابطه فوق را می توان در فرم ماتریس بصورت زیر نوشت :

$$\min_{\ell^0 \varepsilon R} V^T PV \qquad (6\text{-}٤٣)$$

که در آن P یک ماتریس قطری است و عناصر روی قطر آن همان P_j ها یا وزن های مربوطه می باشند.

$$P = diag(P_1, P_2, ..., P_m) \qquad (6\text{-}٤٤)$$

فرمول های ٤٢-6 و ٤٣-6 حالت کلی دارند و فرمول ٤١-6 یک حالت خاص از فرمول ٤٢-6 می باشد. یعنی فرمول ٤١-6 را نیز می توان بصورت $\sum_{i=1}^{n} P_i v_i^2$ نوشت که در آن $P_i = \frac{1}{n}$ (یعنی وزنهای ثابت برای تمام n مشاهده) می باشد. بنابراین فرم ٤٣-6 را بخاطر حالت عام آن از حالا به بعد مورد استفاده قرار خواهیم داد و همینطور است فرمول ٤٠-6 که یک فرمول کلی برای میانگین می باشد.

ملاحظه کنید که وقتی شرط ٤٣-6 برقرار شد ، اختلافها ، v_i ها ، تبدیل یه باقیمانده می شوند و ما این باقیمانده ها را با $\hat{V}$ نشان می دهیم. به این ترتیب ، معادله ٧-٣ را می توان بصورت زیر نوشت :

$$S_L^2 = \sum_{j=1}^{n} P_j \hat{v}_j^2 \qquad (6\text{-}٤٥)$$

و یا بصورت ماتریسی

$$S_L^2 = \hat{v}^T P \hat{v} \qquad (6\text{-}٤6)$$

نوشت.

در نتیجه می توان اصل کمترین مربعات را بصورت زیر بیان کرد. مقدار $\hat{x}$ که بازای آن فرم کوادراتیک $V^T PV$ می نیمم می شود ، واریانس می نیمم را برای نمونه L تضمین می کند. این خاصیت $\hat{x}$ هیچ ارتباطی به تابع پخش PDF ندارد. اگر نمونه L دارای PDF نرمال و یا هر تابع پخش قرینه باشد ، $\hat{x}$ عبارت خواهد بود از محتملترین برآورد از x و یا برآورد با ماکزیمم احتمال از x .

۳-٤-6 واریانس میانگین نمونه

دیدیم که پیدا کردن میانگین نمونه را می توان بعنوان یک مسئله سرشکنی ساده در نظر گرفت. حال می توان این سؤال را مطرح کرد : ماتریس واریانس کوواریانس نتایج بدست آمده ، برحسب ماتریس واریانس کوواریانس نمونه اصلی کدام است؟ بعبارت دیگر واریانس نتایج بدست آمده (میانگین نمونه) چقدر است؟

سؤال فوق را می توان توسط قانون انتشار کوواریانس ، جواب داد (بخش ۱-۳-6) . میانگین نمونه در بخش قبلی ، بصورت زیر بدست آمد (فرمول ٤۰-6) :

$$\hat{x} = P^T \bar{L}$$

با اعمال قانون کووارایانس (فرمول ۱۵-6) واریانس $\hat{x}$ بدست می آید :

$$\sum \hat{x} = B \sum \bar{L} B^T$$

$$S_{\hat{x}}^2 = P^T \sum \bar{L} P \qquad (6 - ٤۷)$$

در اینجا ماتریس $\sum \bar{L}$ هنوز تعریف نشده است. و تنها چیزی که می دانیم مجموعه $\bar{L} \equiv (\bar{\ell}_1, \bar{\ell}_2, ..., \bar{\ell}_m)$ می باشد که عبارتست از گروه های مشاهدات $\bar{\ell}_i$ها با وزنهای مختلف (احتمال های مشاهده شده) P_i . فرض میکنیم که این مشاهدات غیر همبسته بوده و همچنین فرض می کنیم که هر کدام دارای واریانس $S_{\ell_i}^2$ می باشند. در نتیجه ماتریس $\sum \bar{L}$ بصورت زیر خواهد بود :

$$\sum \bar{L} = diag\left(S_{\ell_1}^2, S_{\ell_2}^2, ..., S_{\ell_m}^2 \right) \qquad (6 - ٤۸)$$

با جایگزینی ٤۸- 6 در فرمول ٤۷-6 خواهیم داشت :

$$S_{\hat{x}}^2 = \sum_{j=1}^{m} P_j^2 S_{\ell_j}^2 \qquad (6 - ٤۹)$$

از طرف دیگر مقدار $\hat{x}$ (میانگین نمونه) را می توان با استفاده از نمونه اصلی $L = (\ell_1, \ell_2, ..., \ell_n)$ یعنی مشاهدات گروه بندی نشده $\ell_i, i = 1,2,...,n$ با احتمال های تجربی (وزنهای) یکسان و مساوی $P_i = \frac{1}{n}$ نیز بدست آورد.

$$\hat{x} = \frac{1}{n} \sum_{j=1}^{n} \ell_i = \frac{1}{n}(\ell_1 + \ell_2 + ... + \ell_n) \qquad (6 - ۵۰)$$

با استفاده از قانون انتشار خطاها (۵۰- 6) واریانس $\hat{x}$ بصورت زیر محاسبه می شود :

$$S_{\hat{x}}^2 = \sum_{i=1}^{n} \left(\frac{\partial x}{\partial \ell_i} \right)^2 S_{\ell_i}^2 = \left(\frac{1}{n} \right)^2 \sum_{i=1}^{n} S_{\ell_i}^2 \qquad (6 - ۵۱)$$

در فرمول فوق ، تمام واریانس های $S_{\ell_i}^2$ را مساوی ، فرض کرده و برابر واریانس نمونه ، یعنی S_L^2 قرار می دهیم :

$$S_L^2 = \frac{1}{n}\sum_{i=1}^{n}(\ell_i - \hat{x})^2 \qquad (6\text{-}٥٢)$$

با فرض فوق ، معادله 6-٥١ بصورت زیر در می آید :

$$S_{\hat{x}}^2 = \left(\frac{1}{n}\right)^2 \left(nS_L^2\right) = \left(\frac{1}{n}\right) S_L^2 \qquad (6\text{-}٥٣)$$

فرمول فوق ، نشان می دهد که واریانس میانگین نمونه برابر است با واریانس نمونه تقسیم بر تعداد کل (n) عناصر نمونه * .

بنابراین ما دو فرمول مختلف 6-٤٩ و 6-٥٣ را برای محاسبه $S_{\hat{x}}^2$ بدست آوردیم. در فرمول اول ، فرض بر این است که هر کدام از مشاهدات (در حقیقت گروه هائی از مشاهدات یکسان) دارای واریانس های مختلف $S_{\ell_j}^2$ می باشد. در فرمول دوم ، فرض بر این است که تمام مشاهدات ، متعلق به یک نمونه L با واریانس S_L^2 می باشند. دو فرمول 6-٤٩ و 6-٥٣ باید منجر به یک مقدار برای $S_{\hat{x}}^2$ شوند. پس دو فرمول فوق را مساوی همدیگر قرار میدهیم.

* برحسب علائم بکار گرفته شده قبلی فرمول 6-٥٣ را بصورت زیر نیز می توان نوشت :

$$S_{\ell_i}^2 = S_L^2 / n$$

$$\sum_{j=1}^{m}(P_j^2 S_{\ell_j}^2) = \frac{1}{n} S_L^2 \qquad (6\text{-}٥٤)$$

با در نظر گرفتن اینکه جمع احتمال های تجربی P_j برابر واحد می باشد ، یعنی $\sum_{j=1}^{m} P_j = 1$ ، طرفین رابطه فوق را می توان بصورت زیر تغییر داد :

$$\sum_{j=1}^{m}\left[P_j\left(P_j S_{\ell_j}^2\right)\right] = \sum_{j=1}^{m} P_j\left(\frac{S_L^2}{n}\right) \qquad (6\text{-}٥٥)$$

از رابطه فوق نتیجه می گیریم که :

$$P_j S_{\ell_j}^2 = \frac{S_L^2}{n} = k \quad \text{و} \quad j = 1,2,\ldots,m \qquad (6\text{-}٥6)$$

که در آن k دارای مقدار ثابت برای یک نمونه بخصوص و برابر وارایانس میانگین آن نمونه می باشد. از فرمول 6-٥6 نتیجه می گیریم که :

$$S^2_{\ell_j} = \frac{k}{P_j} \quad \text{و} \quad j = 1,2,...,m \qquad (6-٥٧)$$

رابطه فوق نشان می دهد که بمنظور رسیدن به نتیجه صحیح از فرمول ٤٩-6 باید فرض کنیم که واریانس هر کدام از مشاهدات ، نسبت عکس با وزن آن مشاهده دارد. نتیجه فوق ، معمولاً بصورت اصل زیر بیان می گردد.

وزن یک مشاهده با واریانس آن ، نسبت عکس دارد.

$$P = \frac{k}{S^2} \qquad (6-٥٨)$$

با استفاده از فرمول ٥٧-6 می توان نوشت :

$$P_1 S_1^2 = P_2 S_2^2 = ... = 1 S_0^2 = k \qquad (6-٥٩)$$

که در آن S_0^2 یک مقدارثابت برای نمونه و معروف به واریانس وزن واحد می باشد. میتوان چنین تعبیر کرد که S_0^2 عبارتست از واریانس یک مشاهده خیالی با وزن واحد. اگر میانگین نمونه یعنی $\overline{\ell}$ را در نظر بگیریم S_0 همان $S_{\overline{\ell}}$ خواهد بود.

از فرمولهای ٤6-6 و ٥٣-6 نتیجه می شود که :

$$S^2_{\hat{x}} = \frac{\hat{V}^T P \hat{V}}{n} \qquad (6-6٠)$$

در بخش های بعدی ، اغلب ، به این فرمول اشاره خواهیم کرد.

باید در نظر گرفت که تمام بحث های فوق متکی به قبول واریانس های $S^2_{\ell_i}$ و $S^2_{\overline{\ell}_i}$ میباشد. منظور از معرفی آنها ، بدست آوردن فرمولهای ٥٣-6 و ٥٨-6 که مطابق با بقیه محاسبات سرشکنی هستند می باشد. بطور جدی تر اینست که دو فرمول فوق را بدون اثبات ، یعنی مثل دو تعریف قبول کرد.

٤-٤-6 ماتریس واریانس کوواریانس میانگین یک جند نمونه ای

در بخش قبل دیدیم که میانگین $\overline{\ell}$ یک نمونه دارای واریانس $S^2_{\overline{\ell}}$ مربوط به خودش می باشد. این واریانس به اندازه n برابر کوچکتر از واریانس خود نمونه(S_L^2)می باشد. $S^2_{\overline{\ell}}$ را میتوان بعنوان معیاری از اطمینان روی $\overline{\ell}$ دانست.

واضح است که اطمینان ما روی $\overline{\ell}$ با افزایش تعداد مشاهدات (n) زیادتر خواهد شد.

حال می توان این سؤال را مطرح کرد که آیا میانگین $\overline{L}$ یک چند نمونه ای L نیز دارای ماتریس واریانس کوواریانس مخصوص به خودش است؟ در جواب ، باید گفت بلی. چون مانعی در بین نیست که ما را ازتعریف چنین ماتریسی بازدارد. ماتریس مزبور را می توان با عمومیت دادن حالت یک نمونه ای بصورت زیر تعریف کرد :

$$\sum \overline{L} = \begin{bmatrix} S^2_{\overline{\ell}_1} & S_{\overline{\ell}_1\overline{\ell}_2} & \dots & S_{\overline{\ell}_1\overline{\ell}_S} \\ S_{\overline{\ell}_2\overline{\ell}_1} & S^2_{\overline{\ell}_2} & \dots & S_{\overline{\ell}_2\overline{\ell}_S} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ S_{\overline{\ell}_S\overline{\ell}_1} & S_{\overline{\ell}_S\overline{\ell}_2} & \dots & S^2_{\overline{\ell}_S} \end{bmatrix} \qquad (6\text{-}61)$$

که در آن $S^2_{\overline{\ell}_i} = \frac{1}{n_i} S^2_{\ell_i}$ و $S_{\overline{\ell}_i\overline{\ell}_j} = \frac{1}{n} S_{\ell_i\ell_j}$ می باشند. در اینجا باز نیاز به تساوی عناصر نمونه ها یعنی $n_i = n_j = n$ داریم. این تساوی لازم است در مورد تمام مؤلفه های چند نمونه ای برقرار باشد. یعنی :

$$n_1 = n_2 = \ldots = n_s = n$$

و در نتیجه خواهیم داشت :

$$\sum \overline{L} = \frac{1}{n} \sum L \qquad (6\text{-}62)$$

همینطور ماتریس واریانس کوواریانس که از طریق قانون کوواریانس و از ماتریس واریانس کوواریانس میانگین چند نمونه ای محاسبه می شود ، مربوط به $\hat{x}$ (برآورد آماری از X) می باشد. یعنی :

$$\sum \hat{x} = B \sum \overline{L} B^T \qquad (6\text{-}63)$$

که ماتریس واریانس کوواریانس $\hat{x}$ ، جواب مدل ریلضی زیر می باشد :

$$\hat{x} = F\left(\overline{L}\right)$$

روابط نظیر فوق را می توان برای سایر قوانین انتشار خطاها نیز نوشت. نتایج بدست آمده در این بخش را با جوابهای مثال ۱۴-6 مقایسه کنید.

مثال ۱۵-6

مثالهای ۱-6 و ۳-6 و ٤-6 را در نظر می گیریم. این بار می خواهیم ماتریس واریانس کوواریانس $\sum \hat{x}$ را برای بردار $\hat{x}$ پیدا کنیم.

نخست از فرمول ۱6-6 ماتریس $\sum \bar{L}$ را محاسبه می کنیم :

$$S_{\bar{a}}^2 = \frac{1}{5}S_a^2 = \frac{0.004}{5} = 0.0008cm^2 \quad \text{و} \quad S_{\bar{b}}^2 = \frac{1}{5}S_b^2 = \frac{0.0056}{5} = 0.0011cm^2$$

با در نظر گرفتن $S_{ab} = 0$ خواهیم داشت :

$$\sum \bar{L} = \begin{bmatrix} 0.0008 & 0 \\ 0 & 0.0011 \end{bmatrix} = \frac{1}{5}\sum L$$

حال ماتریس $\sum \hat{x}$ را در معادله ۳6-6 محاسبه می کنیم :

$$\sum \hat{x} = B\left(\frac{1}{5}\sum L\right)B^T = \frac{1}{5}B\sum LB^T = \frac{1}{5}\sum x$$

$$\sum \hat{x} = \begin{bmatrix} 0.00081 & 0.1079 \\ 0.1079 & 21.51254 \end{bmatrix} \Rightarrow \begin{matrix} S_{\hat{a}} = \sqrt{0.00081} = 0.028cm \\ S_{\hat{b}} = \sqrt{21.51254} = 4.64cm \end{matrix}$$

۵-٤-6 روش کمترین مربعات ، ماتریس وزن

روش کمترین مربعات ، بطوریکه در حالت بسیار ساده سرشکنی روی مدل $x = \ell$ اعمال شد(بخش قبل) ، می تواند برای حل مدلهای دیگر نیز تعمیم داده شود. با در نظر گرفتن حالت کلی مدل ریاضی مثل :

$$F(\bar{L} + V, x) = 0$$

که در بخش ۱-٤-6 بیان شد ، می توان در اینجا نیز دنبال چنان مقداری از x بود که مقدار $V^T PV$ (فرم کوادراتیک از اختلافهای وزن دار) را می نیمم کند. یعنی :

$$\min_{\hat{x} \varepsilon R^n} V^T PV \qquad (6\text{-}6٤)$$

اعمال شرط فوق در اکثر مدلهای ریاضی منجر به یک جواب واحد $\hat{x}$ خواهد بود. روش سرشکنی را که از شرط فوق در حل مدلها استفاده می کند روش کمترین مربعات می نامند.

سؤالی که در اینجا باقی می ماند این است که ماتریس P را چگونه باید انتخاب کرد؟ در مورد میانگین نمونه ، بخش قبل ، ماتریس P بصورت زیر بکار رفته بود :

$$P = diag\left({k}/{S_{\ell_1}^2}, {k}/{S_{\ell_2}^2}, \ldots, {k}/{S_{\ell_n}^2} \right)$$

یا

$$P = k\,diag\left({1}/{S_{\ell_1}^2}, {1}/{S_{\ell_2}^2}, \ldots, {1}/{S_{\ell_n}^2} \right)$$

با استفاده از علائمی که در مورد چند نمونه ای بکار گرفته شده ، ماتریس فوق را می توان بصورت زیر نیز نوشت :

$$\boxed{P = k\sum_{L}^{-1}} \qquad (6\text{-}65)$$

رابطه فوق ، نشان می دهد که ماتریس P از ضرب یک مقدار ثابت (k) در عکس ماتریس واریانس کوواریانس میانگین مشاهدات بدست می آید. این ماتریس در حالت ما یک ماتریس قطری است (بنا به فرض اینکه مشاهدات مستقل هستند) .

دوباره ملاحظه می کنیم که از نقطه نظر ریاضی ، اختلاف چندانی بین یک نمونه و یک چند نمونه ای وجود ندارد. در نتیجه می توان با آنها یکسان رفتار کرد. بنابراین فرق عمده بین مسئله ساده سرشکنی (محاسبه میانگین نمونه) و مسئله عمومی سرشکنی وجود ندارد. تنها فرق ، اینست که در حالت " محاسبه میانگین نمونه " بردار $\hat{x}$ فقط دارای یک مؤلفه ، ولی در حالت کلی بردار $\hat{x}$ ممکن است چندین مؤلفه داشته باشد.

تجانس بین دو سرشکنی ساده و حالت کلی این سؤال را پیش می آورد که نقش k (ضریب ثابت که برابر $S_{\hat{x}}^2$ در سرشکنی میانگین نمونه) در متد کمترین مربعات برای حالت کلی (یعنی بردار $\hat{x}$ دارای چندین مؤلفه باشد) چیست؟

در جواب این سؤال ، فقط می گوئیم که ما ماتریس P (در متد کمترین مربعات ، ماتریس وزن نامیده می شود) را از رابطه زیر محاسبه می کنیم :

$$P = k\sum_{L}^{-1} \qquad (6\text{-}66)$$

که در آن K یک عدد ثابت اختیاری است و معنی آن بعداً تشریح خواهد شد. اختیاری بودن k در ماتریس وزن ، هیچ تأثیری در جواب مدل ، یعنی $\hat{x}$ ندارد ، بعداً این مسئله روشن خواهد شد ، چون ضریب k نسبت بین وزن ها یا واریانس های مشاهدات را بهم نمی زند.

در این درس ، ما فقط با دو تیپ مدل ریاضی که مکرر در عمل با آن مواجه می شویم سر و کار خواهیم داشت. در این مدلها ما علائم زیر را بکار خواهیم برد :

n - تعداد مؤلفه های چند نمونه ای اولیه یا اصلی L یا تعداد مشاهدات

u - تعداد مؤلفه های چند نمونه ای بدست آمده x یا تعداد پارامترهای مجهول

r – تعداد معادلات (روابط) مستقل از هم که می تواند بین L و x برقرار گردند.

بعلاوه مدلهای خطی را در نظر خواهیم گرفت.

اولین مدل بصورت زیر می باشد :

$$Ax = L \qquad (6\text{-}67)$$

که در آن A یک ماتریس با n سطر و u ستون ، x یک بردار با u مؤلفه و L یک بردار با n مؤلفه می باشد (n=r>u) . سرشکنی چنین مدلی را سرشکنی پارامتریک (Parammetric Adjustment) و یا سرشکنی معادلات مشاهدات و یا سرشکنی مشاهدات غیر مستقیم و غیره می نامند.

مدل دوم بصورت :

$$BL = C \qquad (6\text{-}68)$$

می باشد که در آن B یک ماتریس با r سطر و n ستون و L یک بردار با n مؤلفه و C یک بردار با r مؤلفه می باشد (r<n) . سرشکنی چنین مدلی را سرشکنی معادلات شرط (Conditional Adjustment) می نامند.

دو مدل فوق ، از آنجائیکه هر دو خطی هستند ، حالات بسیار خاص می باشند. خوشبختانه خیلی از مسائل در عمل ، هرچند که طبیعتاً غیر خطی هستند ، می توانند به خطی تبدیل شوند. بدین علت است که دو مدل فوق مهم می باشند.

6-٤-6 سرشکنی پارامتریک (Parammetric Adjustment)

در این بخش ، ما از سرشکنی مدل خطی 6-6٧ صحبت خواهیم کرد.

$$Ax = L \qquad n > u \qquad (6\text{-}69)$$

برای سرشکنی مدل فوق ، آنرا بصورت زیر تغییر می دهیم :

$$Ax - \left(\overline{L} + V\right) = 0$$

و یا

$$\boxed{V = Ax - \overline{L}} \qquad * \qquad (6\text{-}70)$$

در اینجا A ماتریس ضرایب و x بردار پارامترهای مجهول و $\overline{L}$ بردار مشاهدات یا هر نوع کمیتهای معلوم و ثابت دیگر و V بردار اختلافها که آنها نیز مجهول می باشند.

مدل 6-٧٠ را یک سری معادلات مشاهدات می نامند.

ما دنبال چنان $x = \hat{x}$ هستیم که فرم کوادراتیک $V^T PV$ را می نیمم کند. P عبارتست از ماتریس وزن که برای مشاهدات $\overline{L}$ در نظر گرفته می شود. فرم کوادراتیک فوق که بعضی اوقات فرم کوادراتیک از اختلافهای وزن دار نامیده می شود با در نظر گرفتن معادله 6-٧٠ بصورت زیر نوشته می شود :

$$V^T PV = \left(AX - \overline{L}\right)^T P\left(AX - \overline{L}\right) = \left[(AX)^T - \overline{L}^T\right]\left[PAX - P\overline{L}\right]$$

$$= X^T A^T PAX - \overline{L}^T PAX - X^T A^T P\overline{L} + \overline{L}^T P\overline{L} \qquad (6-٧١)$$

از معادله 66-6 داریم $P = k\sum_{\overline{L}}^{-1}$ که k عبارتست از یک ضریب ثابت و $\sum\overline{L}$ ماتریس واریانس کوواریانس بردار $\overline{L}$ می باشد. ا زآنجائیکه $\sum\overline{L}$ یک ماتریس قرینه است ، ماتریس وزن نیز قرینه بوده و خواهیم داشت $P^T = P$. بنابراین می توان نوشت :

$$\overline{L}^T PAX = X^T A^T P\overline{L} \qquad (6-٧٢)$$

حاصلضرب ماتریس های فوق یک عدد است. بنابراین ترانسپوز آن با خودش مساوی خواهد بود. با جایگزینی ٧٢-6 در ٧١-6 خواهیم داشت :

$$V^T PV = X^T A^T P\overline{L} - 2L^T PAX + \overline{L}^T P\overline{L} \qquad (6-٧٣)$$

* اگر یک مدل غیر خطی $L = F(X)$ داشته باشیم ، می توانیم از بسط توابع به سری تیلور استفاده کرده آنرا خطی کنیم : $L = F\left(X^0\right) + \frac{\partial F}{\partial X}\Big|_{X = X^0}\left(X - X^0\right) + \ldots$

که در آن از ترمهای با درجات بالاتر صرفنظر شده است. با قرار دادن $\Delta X = X - X^0$ و $\Delta L = L - F\left(X^0\right)$ و $A = \frac{\partial F}{\partial X}\Big|_{X = X^0}$ خواهیم داشت :

$$\Delta L = A\Delta X$$

مدل فوق همان فرم ٦٩-6 می باشد. در این حالت نتیجه سرشکنی عبارتست از تصحیح ΔX نسبت به مقدار تقریبی X^0 به خود بردار X .

تابع کوادراتیک ٧٣-6 را که تابعی از x می باشد ، باید نسبت به x می نیمم کرد. اینکار با مساوی قرار دادن مشتقات تابع مزبور نسبت به مؤلفه های بردار x انجام می گیرد :

$$\frac{\partial}{\partial x^i} V^T PV = 0 \qquad i = 1,2,\ldots,u \qquad (6-٧٤)$$

از تابع ٧٣-6 نسبت به بردار x مشتق گرفته مساوی صفر قرار می دهیم :

$$\frac{\partial}{\partial x} V^T PV = 2\hat{x}^T A^T PA - 2\overline{L}^T PA = 0 \qquad *$$

و یا

$$\hat{x}^T A^T PA = \overline{L}^T PA$$

طرفین تساوی فوق را ترانسپوز می کنیم :

$$\boxed{\left(A^T PA\right)\hat{X} = A^T P\overline{L}} \qquad + \qquad (6-٧٥)$$

سیستم معادلات خطی فوق را معادلات نرمال می نامند که معمولاً بصورت اختصاری زیر نوشته می شود :

$$N\hat{X} = U \qquad (6\text{-}76)$$

جائیکه $N = A^T PA$ را ماتریس ضرایب معادلات نرمال یا بطور ساده ماتریس معادلات نرمال و $U = A^T P\overline{L}$ را بردار مقادیر ثابت معادلات نرمال می نامند.

سیستم معادلات نرمال ۷6-6 یک جواب $\hat{X}$ بصورت زیر دارد :

$$\boxed{\hat{X} = N^{-1}U = \left(A^T PA\right)^{-1} A^T P\overline{L}} \qquad (6\text{-}77)$$

به شرطی که ماتریس معادلات نرمال یعنی $N = A^T PA$ دارای ماتریس معکوس باشد.

* از جبر ماتریس ، می دانیم که اگر M یک ماتریس ثابت و قرینه باشد و x یک بردار ، داریم :

$$\frac{\partial}{\partial X} MX = M \qquad \text{و} \qquad \frac{\partial}{\partial X}\left(X^T MX\right) = 2X^T M$$

*ملاحظه کنیدکه معادلات نرمال رامی توان مستقیماًاز ضرب طرفین مدل با $A^T P$ بدست آورد.

ملاحظه کنید که N یک ماتریس قرینه و همیشه مثبت می باشد. *

برای مطالعه تأثیر ماتریس وزن P روی جواب $(\hat{X})$ یک ماتریس وزن دیگر ، P' بصورت زیر در نظر می گیریم :

$$P' = \gamma P \qquad (6\text{-}78)$$

که در آن γ یک ضریب اختیاری می باشد. با قرار دادن P' در ۷۷-6 خواهیم داشت :

$$\hat{X} = \left(A^T P'A\right)^{-1} A^T P'\overline{L} = \left(A^T \gamma PA\right)^{-1} A^T \gamma P\overline{L}$$
$$= \frac{1}{\gamma}\left(A^T PA\right)^{-1} \gamma\left(A^T P'\overline{L}\right) = \left(A^T PA\right)^{-1} A^T P\overline{L} = \hat{X} \qquad (6\text{-}79)$$

نتیجه فوق نشان می دهد که فاکتور k در معادله 66-6 برای محاسبه ماتریس وزن از ماتریس $\sum \overline{L}$ می تواند اختیاری باشد ، بدون اینکه تأثیری در جواب $\hat{X}$ داشته باشد.

باید گفت که بردار اختلافهای V که در فرمول ۷۰-6 تعریف شد بعد از می نیمم کردن ، تبدیل به بردار باقیمانده ها می شوند (بخش ۸-٤) . بردار باقیمانده ها را باسید با حرف دیگری نظیر R نشان داد. چون بردار باقیمانده ها دیگر یک تابع از x نیست ، بلکه یک بردار از کمیتهای ثابت است. بعضی از مؤلف ها علامت $\hat{V}$ را برای باقیمانده ها بکار میبرند و ما هم از آنها پیروی می کنیم (بخش ۲-٤-6 را ملاحظه کنید) . مقادیر $\hat{V}_i$ها مستقیماً از فرمول ۷۰-6 محاسبه می شوند و بعد فیزیکی آنها همان بعد فیزیکی $\overline{L}$ می باشد. سپس مشاهدات سرشکن شده از فرمول $\hat{L} = \overline{L} + \hat{V}$ محاسبه می شوند.

باید در نظر داشت که یکی از خصوصیات اصلی سرشکنی پارامتریک اینست که جواب $\hat{X}$ مستقیماً از فرمول ۷۷-6 بدست می آید.

---

* ماتریس N را یک ماتریس همیشه مثبت گویند ، اگر مقدار فرم کوادراتیک $Y^T NY$ بازای هر بردار اختیاری Y (با ابعاد مناسب ماتریس N) مثبت باشد.

در این مرحله ارزش دارد که دوباره برگردیم به مسئله سرشکنی ساده (محاسبه میانگین یک نمونه) و آنرا مورد بررسی قرار دهیم. بنا به معادله ۷۹-6 می توانیم وزن هر مشاهده را متناسب معکوس با واریانس مربوطه با یک ضریب تناسب اختیاری k اختیار کنیم. اینکار نشان می دهد که وزن ها لازم نیست مساوی احتمال های تجربی که در مجموع برابر واحد بودند $\left(\sum P_i = 1\right)$ اختیار گردند. در این حالت معادلات مشاهدات عبارت خواهند بود از :

$\hat{x} = \bar{\ell}_1 + \hat{V}_1$ با وزن P_1

$\hat{x} = \bar{\ell}_2 + \hat{V}_2$ با وزن P_2

.

.

.

$\hat{x} = \bar{\ell}_n + \hat{V}_n$ با وزن P_n

و یا بصورت ماتریسی

$$A\hat{X} = \bar{L} + \hat{V}$$

که در آن

$$A = \begin{bmatrix} 1 \\ 1 \\ . \\ . \\ . \\ 1 \end{bmatrix} \quad \text{و} \quad \bar{L} = \begin{bmatrix} \bar{\ell}_1 \\ \bar{\ell}_2 \\ . \\ . \\ . \\ \bar{\ell}_n \end{bmatrix}$$

و ماتریس P یک ماتریس قطری بصورت زیر می باشد :

$$P = diag(P_1, P_2, ..., P_n)$$

با قرار دادن ماتریس های فوق در رابطه ۷۷-6 جواب زیر بدست می آید :

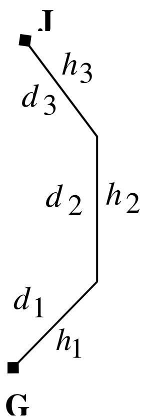
$$\hat{x} = \frac{\sum_{i=1}^{n} P_i \bar{\ell}_i}{\sum_{i=1}^{n} P_i} \qquad (6\text{-}80)$$

که همان میانگین وزن دار نمونه می باشد و با نتیجه بدست آمده در بخش ۲-٤-6 یکی است. با دانستن اینکه $\sum P_i$ در آنجا مساوی واحد بود. فرمول ۸۰-6 یک فرمول عمومی برای محاسبه میانگین یک عده مشاهده وزن دار می باشد.

مثال ۱6-6

یک خط ترازیابی که دو نقطه J و G با ارتفاعات معلوم را به هم متصل می کند ، مورد نظر است. خط ترازیابی به سه نقطه با طولهای d_3, d_2, d_1 تقسیم شده است. اختلاف ارتفاعات h_3, h_2, h_1 مشاهده شده و دارای میانگین های $\bar{h}_3, \bar{h}_2, \bar{h}_1$ می باشد. مشاهدات مزبور مستقل از همدیگر و با واریانس متناسب با طولهای بر قطعه فرض شده است.

مطلوبست تعیین ارتفاعات سرشکن شده نقاط ۱ و ۲ ، یعنی H_2, H_1 با استفاده از سرشکنی پارامتریک.



(شکل 6-6)

از معلومات مسئله داریم :

$n = 3$ تعداد مشاهدات

$u = 2$ تعداد مجهولات

بنابراین ما یک مشاهده اضافی داریم. روابط مستقل بین مشاهدات و مجهولات بصورت زیر نوشته می شود :

$$h_1 = H_1 - H_G$$
$$h_2 = H_2 - H_1$$
$$h_3 = H_J - H_2$$

روابط فوق را می توان در قالب عمومی

$$\underset{3\times 2}{A}\ \underset{2\times 1}{X} = \underset{3\times 1}{L}$$

نوشت که در آن $X = (H_1, H_2)$ و

$$H_1 = h_1 + H_G = L_1$$
$$-H_1 + H_2 = h_2 = L_2$$
$$-H_2 = h_3 - H_J = L_3$$

که در فرم ماتریسی

$$\begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} = \begin{bmatrix} h_1 + H_G \\ h_2 \\ h_3 - H_J \end{bmatrix}$$

نوشته می شود.

معادلات مشاهدات مربوطه بصورت زیر می باشد :

$$H_1 = H_G + (\bar{h}_1 + V_1)$$
$$-H_1 + H_2 = (\bar{h}_2 + V_2)$$
$$-H_2 = -H_J + (\bar{h}_3 + V_3)$$

معادلات فوق را بصورت ماتریسی زیر می توان نوشت :

$$\underset{3\times1}{V} = \underset{3\times2}{A}\ \underset{2\times1}{X} - \underset{3\times1}{\bar{L}}$$

که در آن

$$\underset{3\times1}{V} = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} \quad \text{و} \quad \underset{2\times1}{X} = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} \quad \text{و} \quad \bar{L} = \begin{bmatrix} \bar{h}_1 + H_G \\ \bar{h}_2 \\ \bar{h}_3 - H_J \end{bmatrix} \quad \text{و} \quad \underset{3\times2}{A} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}$$

فرض بر این بوده است که مشاهدات $\bar{h}_3, \bar{h}_2, \bar{h}_1$ غیر همبسته هستند. علاوه بر آن فرض میکنیم که H_G, H_J بدون خطا می باشند. در نتیجه داریم :

$$\sum \bar{L} = diag\left(S^2_{\bar{h}_1}, S^2_{\bar{h}_2}, S^2_{\bar{h}_3} \right)$$

اما $S^2_{\bar{h}_i}$ها متناسب با d_iها می باشند. پس :

$$\sum L = k\,diag(d_1, d_2, d_3)$$

می توان ضریب تناسب k را مساوی واحد انتخاب کرد. بنابراین ماتریس وزن

$$P = k \sum \frac{-1}{L} = diag\left(\frac{1}{d_1}, \frac{1}{d_2}, \frac{1}{d_3} \right)$$

با اعمال روش کمترین مربعات ، معادلات نرمال بشکل زیر خواهند بود :

$$\underset{2\times2}{N}\ \underset{2\times1}{\hat{X}} = \underset{2\times1}{U}$$

که در آن

$$N = A^T P A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{d_1} & 0 & 0 \\ 0 & \frac{1}{d_2} & 0 \\ 0 & 0 & \frac{1}{d_3} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}$$

$$\underset{2\times2}{N}=\begin{bmatrix}\left(\frac{1}{d_1}+\frac{1}{d_2}\right) & -\frac{1}{d_2}\\ -\frac{1}{d_2} & \left(\frac{1}{d_2}+\frac{1}{d_3}\right)\end{bmatrix}$$

$$U=A^T P\bar{L}=\begin{bmatrix}1 & -1 & 0\\ 0 & 1 & -1\end{bmatrix}\begin{bmatrix}\frac{1}{d_1} & 0 & 0\\ 0 & \frac{1}{d_2} & 0\\ 0 & 0 & \frac{1}{d_3}\end{bmatrix}\begin{bmatrix}\bar{h}_1+H_G\\ \bar{h}_2\\ \bar{h}_3-H_J\end{bmatrix}=\begin{bmatrix}\left(\frac{\bar{h}_1+H_G}{d_1}-\frac{\bar{h}_2}{d_2}\right)\\ \left(\frac{\bar{h}_2}{d_2}-\frac{\bar{h}_3-H_J}{d_3}\right)\end{bmatrix}$$

$$\underset{2\times1}{\hat{X}}=\underset{2\times2}{N^{-1}}\ \underset{2\times1}{U}$$

$$N^{-1}=\frac{d_1 d_2 d_3}{d_1+d_2+d_3}\begin{bmatrix}\left(\frac{1}{d_2}+\frac{1}{d_3}\right) & \frac{1}{d_2}\\ \frac{1}{d_2} & \left(\frac{1}{d_1}+\frac{1}{d_2}\right)\end{bmatrix}$$

بعد از انجام ضرب $N^{-1}U$ مجهولات $\hat{X}=\left(\hat{H}_1,\hat{H}_2\right)$ بصورت زیر محاسبه می شوند :

$$\hat{H}_1=H_G+\bar{h}_1+\frac{d_1}{\sum_i d_i}\left(H_J-H_G-\sum_i \bar{h}_i\right)$$

$$\hat{H}_2=H_J-\bar{h}_3-\frac{d_3}{\sum_i d_i}\left(H_J-H_G-\sum_i \bar{h}_i\right)$$

حال می توانیم باقیمانده های $\hat{V}_i$ را بصورت زیر محاسبه کنیم :

$$\hat{V}=A\hat{X}-\bar{L}$$

$$\hat{V}_i=\frac{H_J-H_G-\sum \bar{h}_i}{\sum_i d_i}d_i \qquad \text{و} \qquad i=1,2,3$$

در خاتمه مشاهدات سرشکن شده بصورت زیر خواهند بود :

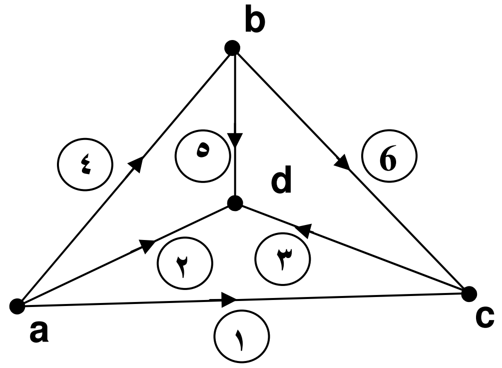
$$\hat{L}=\bar{L}+\hat{V}$$

با در نظر گرفتن اینکه H_G,H_J بدون خطا در نظر گرفته شده اند ، خواهیم داشت :

$$\hat{h}_i=\bar{h}_i+\hat{V}_i \qquad \text{و} \qquad i=1,2,3$$

مثال ۱۷-6

یک شبکه ترازیابی محلی مرکب از شش خط ، شکل ۷-6 ، ترازیابی شده است. جهت فلش ها در شکل نماینده جهت افزایش ارتفاع می باشند. جدول زیر مشاهدات انجام شده بصورت اختلات ارتفاع خط ها همراه با طول هر خط را نشان می دهد.



(شکل ۷-6)

شماره خط	ایستگاه از	ایستگاه به	$\bar{h}_i$ متر	طول ℓ_i کیلومتر
۱	a	c	6·16	٤
۲	a	d	۱۲·٥۷	۲
۳	c	d	6·٤۱	۲
٤	a	b	۱·۰۹	٤
٥	b	d	۱۱·٥۸	۲
6	b	c	٥·۱۷	٤

فرض بر اینست که واریانس های $S_{h_i}^2$ها متناسب با طول ℓ_iها می باشد. ارتفاع نقطه a مساوی صفر فرض شده است. مطلوبست سرشکنی شبکه فوق با روش سرشکنی پارامتریک و تعیین برآوردهای کمترین مربعات $\hat{H}_d ,\hat{H}_c ,\hat{H}_b$ از مجهولات H_d ,H_c ,H_b (ارتفاع نقاط b و c و d).

راه حل :

تعداد مشاهدات مستقل $n = 6$ و تعداد مجهولات $u = 3$ می باشد. در نتیجه ۳ مشاهده اضافی وجود دارد. یعنی ۳ درجه آزادی وجود دارد. مدل ریاضی در اینجا خطی می باشد.

$$\underset{6\times3}{A}\;\underset{3\times1}{X} = \underset{6\times1}{L}$$

که در آن $X = \left(H_b, H_c, H_d\right)$ می باشد. تعداد شش معادله مشاهده وجود دارد که بصورت زیر می باشند :

$$\bar{h}_1 + V_1 = H_c - H_a = H_c - 0.0 = H_c$$
$$\bar{h}_2 + V_2 = H_d - H_a = H_d$$
$$\bar{h}_3 + V_3 = H_d - H_c$$
$$\bar{h}_4 + V_4 = H_b - H_a = H_b$$
$$\bar{h}_5 + V_5 = H_d - H_b$$
$$\bar{h}_6 + V_6 = H_c - H_b$$

معادلات مشاهدات ، بعد از جایگزینی مشاهدات بصورت زیر در می آیند :

$$V_1 = H_c - 6.16$$
$$V_2 = H_d - 12.57$$
$$V_3 = -H_c + H_d - 6.41$$
$$V_4 = H_b - 1.09$$
$$V_5 = -H_b + H_d - 11.58$$
$$V_6 = -H_b + H_c - 5.07$$

معادلات به فرم ماتریسی زیر می توانند نوشته شوند :

$$\underset{6\times1}{V} = \underset{6\times3}{A}\;\underset{3\times1}{X} - \underset{6\times1}{\bar{L}}$$

که در آن

$$\underset{6\times1}{V} = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \\ V_6 \end{bmatrix} \quad \text{و} \quad \underset{3\times1}{X} = \begin{bmatrix} H_b \\ H_c \\ H_d \end{bmatrix} \quad \text{و} \quad \underset{6\times1}{\bar{L}} = \begin{bmatrix} 6.16 \\ 12.57 \\ 6.41 \\ 1.09 \\ 11.58 \\ 5.07 \end{bmatrix}$$

و ماتریس ضرایب A به شکل :

$$\underset{6\times3}{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

از آنجائیکه ما هیچ اطلاعی در مورد همبستگی اختلاف ارتفاعات مشاهده شده $\overline{h_i}$ نداریم آنها را غیر همبسته فرض می کنیم. در نتیجه ماتریس واریانس کوواریانس مشاهدات بصورت :

$$\sum \overline{L} = diag(4,2,2,4,2,4)$$

خواهد بود. در ماتریس فوق ، مقدار k (ضریب ثابت) مساوی واحد قرار داده شده است.

ماتریس وزن مربوط به ماتریس فوق ، بص.رت زیر می باشد :

$$\underset{6\times6}{P} = diag(0.25,0.5,0.5,0.25,05,0.25)$$

معادلات نرمال ، بصورت :

$$\underset{3\times3}{N}\ \underset{3\times1}{\hat{X}} = \underset{3\times1}{U}$$

بوده و جواب آن بشکل $\hat{X} = N^{-1}U$ خواهد بود.

$$N = A^T PA = \begin{bmatrix} 0 & 0 & 0 & 1 & -1 & 1 \\ 1 & 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.25 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.25 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1.00 & -0.25 & -0.5 \\ -0.25 & 1.00 & -0.5 \\ -0.5 & -0.5 & 1.5 \end{bmatrix}$$

ملاحظه کنید که ماتریس N یک ماتریس قرینه مثبت معین می باشد.

$$\underset{3\times3}{N^{-1}} = \begin{bmatrix} 1.6 & 0.8 & 0.8 \\ 0.8 & 1.6 & 0.8 \\ 0.8 & 0.8 & 1.2 \end{bmatrix}$$

$$\underset{3\times1}{U} = \begin{bmatrix} 0 & 0 & 0 & 0.25 & -0.5 & -0.25 \\ 0.25 & 0 & -0.5 & 0 & 0 & 0.25 \\ 0 & 0.5 & 0.5 & 0 & 0.5 & 0 \end{bmatrix} \begin{bmatrix} 6.16 \\ 12.57 \\ 6.41 \\ 1.09 \\ 11.58 \\ 5.07 \end{bmatrix} = \begin{bmatrix} -6.7850 \\ -0.3975 \\ 15.2800 \end{bmatrix}$$

$$\hat{X} = N^{-1}U = \begin{bmatrix} 1.6 & 0.8 & 0.8 \\ 0.8 & 1.6 & 0.8 \\ 0.8 & 0.8 & 1.2 \end{bmatrix} \begin{bmatrix} -6.7850 \\ -0.3975 \\ 15.2800 \end{bmatrix} = \begin{bmatrix} 1.05 \\ 6.16 \\ 12.59 \end{bmatrix}$$

بنابراین ارتفاعات مجهول بصورت زیر بدست آمده اند :

$\hat{H}_B = 1.05$ متر و $\hat{H}_c = 6.16$ متر و $\hat{H}_d = 12.59$ متر

با جایگزینی $\hat{X}$ در فرمول مربوطه مقدار باقیمانده بدست می آید :

$$\hat{V} = A\hat{X} - \overline{L} \qquad \hat{V} = \begin{bmatrix} 0.00\ \text{متر} \\ 0.02\ \text{متر} \\ 0.02\ \text{متر} \\ -0.04\ \text{متر} \\ -0.04\ \text{متر} \\ 0.04\ \text{متر} \end{bmatrix}$$

مشاهدات سرشکن شده $\hat{h}_i$ از فرمول زیر محاسبه می شوند :

$$\hat{h}_i = \overline{h}_i + \hat{V}_i \qquad \text{و} \qquad i = 1,2,\ldots,6$$

$$\hat{h} = \begin{bmatrix} 6.16 \\ 12.57 \\ 6.41 \\ 1.09 \\ 11.58 \\ 5.07 \end{bmatrix} + \begin{bmatrix} 0.00 \\ 0.02 \\ 0.02 \\ -0.04 \\ -0.04 \\ 0.04 \end{bmatrix} = \begin{bmatrix} 6.16 \\ 12.59 \\ 6.43 \\ 1.05 \\ 11.54 \\ 5.11 \end{bmatrix} \text{متر}$$

محاسبات سرشکنی می تواند از طریق محاسبه ارتفاعات نقاط b و c و d بوسیله مشاهدات سرشکن شده و مقایسه آنها با جواب های بدست آمده کنترل گردد.

٧-٤-6 ماتریس واریانس کوواریانس بردار $\hat{X}$ در سرشکنی پارامتریک ، فاکتور واریانس و ماتریس ضریب وزن

در سرشکنی پارامتریک جواب مدل $AX = L$ بصورت معادله ٧٧-6 ، یعنی :

$$\hat{X} = \left(A^T PA\right)^{-1} A^T P\overline{L} = N^{-1} A^T P\overline{L}$$

داده شده است. رابطه فوق می تواند بفرم اختصاری

$$\hat{X} = B\bar{L} \quad \text{که در آن} \tag{6-81}$$

$$B = N^{-1}A^T P \tag{6-82}$$

می باشد نوشته شود. ماتریس واریانس کوواریانس بردار $\hat{X}$ ، یعنی $\sum_{\hat{X}}$ می تواند بطور ساده از اعمال قانون کوواریانس (معادله ۱۵-6) روی ۸۱-6 بدست آید.

$$\sum_{\hat{X}} = B\sum_{\bar{L}} B^T \tag{6-83}$$

از معادله 66-6 داریم :

$$P = k\sum_{\bar{L}}^{-1}$$

با معکوس کردن طرفین رابطه فوق ، خواهیم داشت :

$$\sum_{\bar{L}} = kP^{-1} \tag{6-84}$$

با جایگزینی ۸۲-6 و ۸٤-6 در معادله ۸۳-6 خواهیم داشت :

$$\sum_{\hat{X}} = \left(N^{-1}A^T P\right)kP^{-1}\left(N^{-1}A^T P\right)^T \tag{6-85}$$

هر دو ماتریس P و N قرینه هستند و ترانسپوز آنها با خودشان برابرند و همچنین

$$\left(N^{-1}\right)^T = N^{-1}$$

$$\sum_{\hat{X}} = kN^{-1}A^T PP^{-1}PAN^{-1} = kN^{-1}A^T PAN^{-1} = kN^{-1}NN^{-1} = kN^{-1}$$

یعنی داریم :

$$\boxed{\sum_{\hat{X}} = kN^{-1} = k\left(A^T PA\right)^{-1}} \tag{6-86}$$

از طرف دیگر با قرار دادن $P = k\sum_{\bar{L}}^{-1}$ در معادله ۸6-6 خواهیم داشت :

$$\sum_{\hat{X}} = k\frac{1}{k}\left(A^T\sum_{\bar{L}}^{-1}A\right)^{-1} = \left(A^T\sum_{\bar{L}}^{-1}A\right)^{-1} \tag{6-87}$$

رابطه فوق نشان می دهد که ماتریس واریانس کوواریانس مجهولات ، بستگی به ضریب k ندارد. در حقیقت ، عدم بستگی فوق ، موقعی درست است که ما عناصر ماتریس $\sum_{\bar{L}}$ را بطور صحیح داشته باشیم. بدبختانه اغلب ، ماتریس $\sum_{\bar{L}}$ تا ضریب k مشخص است. یعنی ما فقط واریانس ها و کوواریانس های نسبی بین مشاهدات را می دانیم. بدین معنی که ما مجبوریم بدون ازاطلاع از مقدار واقعی ضریب k ، با ماتریس وزن $k\sum_{\bar{L}}^{-1}$ کار کنیم. بنابراین ماتریس $\sum_{\hat{X}}$ را نمی توان از رابطه ۸۷-6 بدست آورد.

اگر ما فرم کوادراتیک * $\hat{V}^T P\hat{V}$ را بسط بدهیم و فرض کنیم که مشاهدات L فقط تحت تأثیر خطاهای اتفاقی هستند ، یک برآورد برای ضریب k ، یعنی $\hat{k}$ محاسبه نمی شود.

$$\hat{V}^T P\hat{V} = (n-u)\hat{k} \qquad (6-۸۸)$$

ضریب $(n-u)$ عبارتست از اختلاف بین تعداد مشاهدات مستقل n و تعداد پارامترهای مجهول u ، یا بعبارتی تعداد مشاهدات اضافی که بعضی اوقات با df نشان داده می شود. یعنی درجات آزادی (degree of freedom)

$$df = n-u \qquad (6-۸۹)$$

---

* بردار $\hat{V}$ عبارتست از بردار باقیمانده هائی که از طریق سرشکنی به طریقه کمترین مربعات بدست آمده است.

درجه آزادی ، یعنی df باید بزرگتر از صفر باشد تا بتوان سرشکنی انجام داد. بنابراین معادله 6-۸۸ بصورت زیر خواهد بود :

$$\boxed{\hat{k} = \frac{\hat{V}^T P\hat{V}}{df}} \qquad (6-۹۰)$$

k را معمولاً فاکتور واریانس اولیه و $\hat{k}$ را فاکتور واریانس برآورد شده و یا فاکتور واریانس ثانوی می نامند. حال با استفاده از مقدار برآورد شده $\hat{k}$ بجای مقدار اولیه k میتوان مقدار برآورد شده از ماتریس $\sum_{\hat{X}}$ ، یعنی $\hat{\sum}_{\hat{X}}$ را بدست آورد.

$$\boxed{\hat{\sum}_{\hat{X}} = \hat{k}N^{-1} = \hat{k}\left(A^T PA\right)^{-1} = \frac{\hat{V}^T P\hat{V}}{df}\left(A^T PA\right)^{-1}} \qquad (6-۹۱)$$

ماتریس $\hat{\sum}_{\hat{X}}$ معروف به ماتریس برآورد شده واریانس کوواریانس بردار $\hat{X}$ می باشد.

برای مطالعه تأثیر فاکتور واریانس انتخابی k روی ماتریس $\hat{\sum}_{\hat{X}}$ ، ما ماتریس وزن $P = k\sum_L^{-1}$ را با ضریب دیگری یعنی $P' = k'\sum_L^{-1} = \gamma P$ انتخاب کرده در رابطه 6-۹۱ قرار می دهیم :

$$\hat{\sum}_{\hat{X}} = \frac{\hat{V}^T P'\hat{V}}{df}\left(A^T P'A\right)^{-1} = \gamma\frac{\hat{V}^T P\hat{V}}{df}\frac{1}{\gamma}\left(A^T PA\right)^{-1} = \frac{\hat{V}^T P\hat{V}}{df}\left(A^T PA\right)^{-1} = \hat{\sum}_{\hat{X}}$$

تساوی فوق نشان می دهد که ماتریس $\hat{\sum}_{\hat{X}}$ مستقل از انتخاب فاکتور واریانس k می باشد.

اغلب در محاسبات سرشکنی اتفاق می افتد که ما با پارامترهای بدست آمده $\hat{X}$ را بعنوان " مشاهدات " در مرحله سرشکنی بعدی در نظر می گیریم. در این حالت ما باید وزن های واسط را نیز در نظر بگیریم. می دانیم که ماتریس وزن ، یک بردار مشاهدات نسبت عکس با ماتریس واریانس کوواریانس آن دارد (معادله 66-6) . بنابراین می توان ماتریس معادلات نرمال N را بلافاصله بعنوان ماتریس وزن بردار $\hat{X}$ بکار برد ، به معادله ۸6-6 رجوع شود. بنابراین ماتریس N را *ماتریس ضریب وزن* نیز می گویند و جذر عناصر روی قطر اصلی آن را ضرایب وزن می نامند.

ملاحظه کنید که بردار $\hat{X}$ غیر همبسته خواهد بود ، اگر N^{-1} یک ماتریس قطری باشد یعنی وقتیکه N قطری است. در این حالت هر معادله از سری معادلات نرمال می تواند بطور مستقل حل شود. بطوریکه از حل هر معادله یکی از مؤلفه های $\hat{X}$ پیدا خواهد شد. همبستگی بردار $\hat{X}$ ارتباط دوری با همبستگی بردار $\bar{L}$ دارد. بردار $\hat{X}$ غیر همبسته خواهد بود ، اگر بردار $\bar{L}$ غیر همبسته باشد (یعنی ماتریس P قطری باشد) و همچنین ماتریس A متعامد (Orthogonal) باشد. از طرف دیگر ماتریس N ممکن است حتی بازای بعضی دیگر از حالتهای کلی ماتریس های P و A قطری گردد.

یکبار دیگر برمی گردیم به مسئله سرشکنی ساده (تعیین میانگین نمونه) ، بخش ۳-٤-6. در این حالت ، معادالت نرمال تبدیل به یک معادله نرمال می شوند. معادله ٤٠-6 و از طرف دیگر ، با استفاده از معادله ۹۱-6 واریانس برآورد شده میانگین $\hat{X}$ بصورت زیر خواهد بود:

$$\hat{S}_{\hat{x}}^2=\frac{\hat{V}^T P\hat{V}}{df}=\frac{\hat{V}^T P\hat{V}}{n-1} \qquad (6-۹۲)$$

بدیهی است که فرق واریانس برآورد شده فوق با واریانس $S_{\bar{x}}^2$ (معادله ٦٠-6) فقط در مخرج آنها می باشد. مشابه فرمول فوق یک کمیت آماری جدید به نام <u>واریانس برآورد شده نمونه L</u> را بصورت زیر تعریف می کنیم :

$$\boxed{S_L^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(\ell_i-\bar{\ell}\right)^2=\frac{1}{n-1}\sum_{i=1}^{n}\hat{V}_i^2} \qquad (6-۹۳)$$

فرمول فوق را با فرمول ۳-6 مقایسه کنید. فرمول فوق موقعی بکار می رود که میانگین نمونه یعنی $\bar{\ell}$ نیز از طریق فرمول میانگین برآورد شده باشد (یعنی از قبل معلوم نباشد) . بعهده دانشجویان است که نشان بدهند با استفاده از واریانس های برآورد شده مشاهدات گروهی

(بخش ۲-٤-6) و با استدلال بخش ۲-٤-6 و ۳-٤-6 می توان فرمول ۹۲-6 را بجای فرمول ۶۰-6 بدست آورد.

واریانس های برآورد شده نمونه L و میانگین آن $(\bar{\ell})$ را می توان با وجود وزن های غیر نرمال (وزن هائی که برای آنها $\sum_{i=1}^{n} P_i \neq 1$ می باشد) نیز بدست آورد. فرمولهای مربوطه عبارتند از :

$$\hat{S}_L^2 = \frac{1}{n-1}\sum_{i=1}^{n} P_i \hat{V}_i^2 \tag{6-94}$$

$$\hat{S}_{\bar{\ell}}^2 = \frac{1}{(n-1)\sum_{i=1}^{n} P_i}\sum_{i=1}^{n} P_i \hat{V}_i^2 \tag{6-95}$$

در خاتمه این بخش سعی می کنیم معنی فاکتور واریانش را که برا ی اولین بار در بخش ۵-٤-6 عنوان شد ، تفسیر کنیم. فرض کنیم که ماتریس نرمال یک سرشکنی ساده یک ماتریس یکه باشد ، یعنی $(N = I)$. در این حالت ماتریس واریانس کوواریانس بردار جواب $\hat{X}$ برابر

$$\sum \hat{X} = diag(k, k, \ldots, k) \tag{6-96}$$

خواهد بود. ماتریس فوق نشان می دهد که واریانس های مؤلفه های بردار $\hat{x}_i, \hat{X}$ مساوی و برابر k می باشند ($S_{\hat{x}_i}^2 = k$) . از آنجائیکه جذر عناصر قطر اصلی ماتریس N را که همه مساوی واحد هستند می توان بعنوان وزن مؤلفه های $\hat{x}_i$ درنظر گرفت و با در نظر گرفتن تساوی های زیر

$$P_1 S_1^2 = P_2 S_2^2 = \ldots = P_u S_u^2 = k \tag{6-97}$$

که از ماتریس ۹6-6 نتیجه می شود نقش فاکتور واریانس k مشخص می شود. آنرا می توان واریانس وزن واحد ، بخش ۳-٤-6 ، در نظر گرفت و معمولاً آنرا با S_0^2 و یا σ_0^2 (برای حالتیکه واریانس مشخص باشد) نشان می دهند. همینطور $\hat{k}$ را می توان با $\hat{S}_0^2$ و یا $\hat{\sigma}_0^2$ نشان داد.

با قرار دادن علامت $\hat{\sigma}_0^2$ برای $\hat{k}$ و علامت Q بجای ماتریس ضریب وزن پارامترهای مجهول ، یعنی N^{-1} ، معادلات ۹۰-6 و ۹۱-6 بصورت زیر درمی آیند :

$$\hat{\sigma}_0^2 = \frac{\hat{V}^T P \hat{V}}{df} \tag{6-98}$$

$$\hat{\sum}_{\hat{x}} = \hat{\sigma}_0^2 Q \tag{6-۹۹}$$

مثال ۱۸-6

مطلوبست محاسبه ماتریس واریانس کوواریانس برآورد شده ، $\hat{\sum}_{\hat{x}}$ ، پارامترهای سرشکن شده در مثال ۱6-6 .

ماتریس $\hat{\sum}_{\hat{x}}$ از فرمول ۹۹-6 محاسبه می شود. برای اینکار ، نخست معلومات مورد نیاز را از مثال ۱6-6 در زیر می نویسیم :

$$\underset{3\times 3}{\hat{V}^T} = \frac{H_J - H_G - \sum_i \bar{h}_i}{\sum_i d_i}[d_1, d_2, d_3]$$

$$\underset{3\times 3}{P} = diag\left[\frac{1}{d_1}, \frac{1}{d_2}, \frac{1}{d_3}\right] \quad \text{و} \quad df = n - u = 3 - 2 = 1$$

حال برای محاسبه ماتریس $\hat{\sum}_{\hat{x}}$ ضربهای زیر را انجام می دهیم :

$$\hat{V}^T P = \frac{H_J - H_G - \sum_i \bar{h}_i}{\sum_i d_i}[1,1,1]$$

$$\hat{\sigma}_0^2 = \frac{\hat{V}^T P V}{df} = \left(H_J - H_G - \sum_i \bar{h}_i\right)^2 \Big/ \sum_i d_i$$

$$\hat{V}^T P \hat{V} = \left(H_J - H_G - \sum_i \bar{h}_i\right)^2 \Big/ \sum_i d_i$$

$$\underset{2\times 2}{Q} = N^{-1} = \frac{d_1 d_2 d_3}{\sum_i d_i}\begin{bmatrix} \frac{d_2 + d_3}{d_2 d_3} & \frac{1}{d_2} \\ \frac{1}{d_2} & \frac{d_1 + d_2}{d_1 d_2} \end{bmatrix}$$

$$\hat{\sum}_{\hat{x}} = \hat{\sigma}_0^2 Q = \left(\frac{H_J - H_G - \sum_i \bar{h}_i}{\sum_i d_i} \right)^2 \begin{bmatrix} d_1(d_2+d_3) & d_1 d_3 \\ d_1 d_3 & d_3(d_2+d_1) \end{bmatrix}$$

مثال ۱۹-6

مطلوبست تعیین ماتریس واریانس کوواریانس برآورد شده $\hat{\sum}_{\hat{x}}$ پارامترهای $\hat{X}$ در مثال ۱۷-6 .

در اینجا از معادلات ۹۸-6 و ۹۹-6 استفاده می کنیم. از مثال ۱۷-6 داریم :

$$\underset{1\times6}{\hat{V}^T} = [0.00, 0.02, 0.02, -0.04, -0.04, 0.04]m$$

$$\underset{6\times6}{P} = diag(0.25, 0.5, 0.5, 0.25, 0.5, 0.25)$$

$$df = n - u = 6 - 3 = 3$$

$$\hat{V}^T P \hat{V} = 0.002m^2$$

$$\hat{\sigma}_0^2 = \frac{\hat{V}^T P \hat{V}}{df} = \frac{0.002}{3} \cong 0.00067m^2$$

همینطور از مثال ۱۷-6 داریم :

$$\underset{3\times3}{Q} = N^{-1} = \begin{bmatrix} 1.6 & 0.8 & 0.8 \\ 0.8 & 1.6 & 0.8 \\ 0.8 & 0.8 & 1.2 \end{bmatrix}$$ (بدون بعد فیزیکی)

و در خاتمه

$$\hat{\sum}_{\hat{x}} = \hat{\sigma}_0^2 Q = 10^{-4} \begin{bmatrix} 10.67 & 5.33 & 5.33 \\ 5.33 & 10.67 & 5.33 \\ 5.33 & 5.33 & 8.0 \end{bmatrix} m^2 \Rightarrow \hat{\sum}_{\hat{x}} = \begin{bmatrix} 10.67 & 5.33 & 5.33 \\ 5.33 & 10.67 & 5.33 \\ 5.33 & 5.33 & 8.0 \end{bmatrix} cm^2$$

۸-٤-6 بعضی خصوصیات بردار جواب سرشکنی پارامتریک

می توان نشان داد که انتخاب ماتریس P (متناسب با عکس ماتریس واریانس کوواریانس $\sum \bar{L}$) بعنوان ماتریس وزن مشاهدات $\bar{L}$ و انتخاب روش کمترین مربعات (می نیمم کردن $V^T PV$) در سرشکنی یک مدل ، باعث می شود که جواب برآورد شده $\hat{X}$ دارای ماتریس

واریانس کوواریانس $\left(\hat{\sum}\hat{x}\right)$ باشد که مجموع عناصر روی قطر اصلی آن می نیمم است.

بعبارت دیگر انتخاب ماتریس وزن P بصورت $P=\hat{\sigma}_0^2\sum_L^{-1}$ و جستجوی می نیمم مقدار برای $V^T PV$ در سرشکنی پارامتریک ، منجر به جوابی $\left(\hat{X}\right)$ می شود که در عین حال ، شرط زیر را نیز دارا می باشد :

$$\min_{X\varepsilon R^u} trace\sum\hat{X} \quad * \qquad (6-١٠٠)$$

---

* منظور از کلمه trace ، یعنی مجموع عناصر روی قطر اصلی یک ماتریس.

نتیجه فوق ، مشابه نتیجه ای است که از اعمال کمترین مربعات روی چند نمونه ای تصادفی (بخش ٤-٥) بدست آمد و در اینجا فرصت اثبات آن نیست.

همینطور می توان نشان داد که برای یک چند نمونه ای مستقل از مشاهدات $L=\left(L_1,L_2,...,L_n\right)$ که دارای تابع پخش نرمال بصورت

$$\Phi\left(L_0,S,L\right)=\prod_{i=1}^{n}\frac{1}{S_i\sqrt{2X}}e^{-\frac{\left(L^i-L_i^0\right)^2}{2S_i^2}} \qquad (6-١٠١)$$

می باشد. محتملترین برآورد از L_0 مقداری است که تابع $V^T PV$ را می نیمم کند. برای نشان دادن آن می نویسیم :

$$\Phi\left(L_0,S,L\right)=\frac{1}{\left(2X\right)^{n/2}\prod_{i=1}^{n}S_i}e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{V_i^2}{S_i^2}}=\frac{1}{\left(2X\right)^{n/2}\prod_{i=1}^{n}S_i}e^{-\frac{1}{2k}V^T PV}$$

تابع فوق ماکزیمم است ، اگر هر دو تابع $V^T PV$ و مجموع عناصر قطر اصلی ماتریس $\sum\hat{x}$ می نیمم باشند.

٩-٤-6 وزن های نسبی ، اهمیت آماری فاکتورهای واریانس اولیه و ثانوی

در بخش 6-٤-6 دیدیم که انتخاب غاکتور واریانس اولیه σ_0^2 و یا k تأثیری در کمیت برآورد شده $\hat{X}$ ندارد و همینطور در بخش ٧-٤-6 ثابت شد که ماتریس واریانس کوواریانس برآورد شده $\hat{X}$ ، $\sum\hat{x}$ ، مستقل از فاکتور واریانس اولیه σ_0^2 و یا k می باشد. بنابراین برای بدست

آوردن $\hat{X}$ و $\hat{\sum} \hat{X}$ هر سری از وزن های نسبی ، یعنی $P=\hat{\sigma}_0^2 \sum_{L}^{-1}$ با فاکتور واریانس σ_0^2 اختیاری می تواند در نظر گرفته شود. از طرف دیگر ماتریس معادلات نرمال ، یعنی $N=A^T PA$ و فاکتور واریانس برآورد شده ، یعنی $\hat{\sigma}_0^2 = \hat{V}^T P\hat{V} / df$ تحت تأثیر فاکتور اولیه σ_0^2 هستند.

در عمل از خواص فاکتور σ_0^2 به دو منظور زیر استفاده می کنند :

نخست اینکه فاکتور σ_0^2 را چنان انتخاب می کنند که ماتریس نرمال N بدقت قابل معکوس کردن باشد. برای اینکار فاکتور σ_0^2 را چنان انتخاب می کنند که میانگین عناصر داخل ماتریس N نزدیک به واحد باشد.

دوم اینکه از σ_0^2 برای امتحان مدل ریاضی و توافق آن با مشاهدات و امتحان صحت انتخاب ماتریس واریانس کوواریانس مشاهدات ، یعنی $\sum_{L}$ استفاده می کنند.

معمولاً وقتی هیچ اطلاعی از فاکتور σ_0^2 نداریم آنرا مساوی واحد $\left(\sigma_0^2=1\right)$ انتخاب میکنیم. سپس پس از انجام سرشکنی کمترین مربعات مقدار برآورد شده آنرا ، یعنی $\hat{\sigma}_0^2 = \dfrac{\hat{V}^T P\hat{V}}{df}$ بدست می آوریم. نسبت $\dfrac{\hat{\sigma}_0^2}{\sigma_0^2}$ می تواند صحت ماتریس $\sum_{L}$ و توافق مدل ریاضی را مشخص کند. با انتخاب $\sigma_0^2=1$ باید مقدار $\hat{\sigma}_0^2$ هم مساوی واحد بدست آید. اگر چنین نباشد ، ماتریس $\sum L$ مفروض را مورد بررسی قرار داده و مقدار بدست آمده $\hat{\sigma}_0^2$ را بجای σ_0^2 قرار داده ، وزن های جدید را برای مشاهدات محاسبه می کنیم. اگر واریانس ها و کوواریانس های جدید از حدی باشند که تجربه نشان می دهد ، معلوم می شود که ماتریس $\sum_{L}$ قبلی درست انتخاب شده و این بار مدل ریاضی را مورد امتحان قرار داده ، صحت روابط موجود بین مشاهدات و پارامترهای مجهول را بررسی می کنیم.

از روش فوق همچنین در کشف خطاهای سیستماتیک روی مشاهدات استفاده می کنند ، چون مشاهداتی که دارای خطاهای سیستماتیک هستند با مدل ریاضی توافق نخواهند داشت . این عدم توافق در مقدار فرم کوادراتیک $V^T PV$ و در نتیجه روی $\hat{\sigma}_0^2$ متجلی خواهد شد.

رابطه تئوری موجود بین واریانس فاکتور اولیه و ثانوی ما را قادر می سازد که اعتبار فرضیات خود را مورد تست آماری قرار دهیم. در اینجا فرصت پرداختن به جزئیات تست آماری فوق

نیست ، فقط می پردازیم به بررسی سرشکنی خط ترازیابی مورد بحث مثالهای ۱۷-6 و ۱۹-6. در محاسبه ماتریس وزن P فرض کردیم که $\hat{\sigma}_0^2 = 1$ باشد. بعد از سرشکنی مقدار $\hat{\sigma}_0^2 \cong 0.00067$ بدست آمد. نسبت $\hat{\sigma}_0^2 / \sigma_0^2$ برابر 1500 می باشد که خیلی دور از عدد 10 می باشد. نتیجه فوق نشان می دهد که ماتریس واریانس کوواریانس $\sum \overline{L}$ خیلی بدبینانه انتخاب شده است ، یعنی واریانس های واقعی مشاهدات خیلی کمتر از مقادیر انتخاب شده هستند.

6-4-10 سرشکنی معادلات شرط (Conditional Adjustment)

در این بخش ، ما از سرشکنی مدل خطی 68-6 صحبت خواهیم کرد.

$$BL = C \quad \text{و} \quad r < n \qquad (6\text{-}102)$$

مدل فوق تعداد r شرط مستقل و خطی بین n مشاهده L را نشان می دهد. C عبارتست از یم بردار با r مؤلفه ثابت که بوسیله شرایط تعیین می شوند. برای سرشکنی مدل فوق ، آنرا بصورت زیر تغییر می دهیم :

$$B(\overline{L} + V) - C = 0$$

و یا

$$BV + W = 0 \quad * \qquad (6\text{-}103)$$

که در آن

$$\boxed{W = B\overline{L} - C} \qquad (6\text{-}104)$$

سیستم معادلات 3-1-6 به نام معادلات شرط معروف می باشند که در آن B ماتریس ضرایب و V بردار اختلافها و W بردار مقادیر ثابت می باشد. در اینجا نیز n عبارتست از تعداد مشاهدات و r عبارتست از تعداد معادلات مستقل شرط. باید گفت که در معادلات شرط ، هیچکدام از پارامترهای مجهول ، ظاهر نمی شوند. اختلافهای V تنها مجهولات فوق میباشند.

در اینجا نیز هدف ، بدست آوردن آن مقدار $\hat{V}$ از بردار V است که فرم کوادراتیک $V^T PV$ را می نیمم کند. ماتریس وزن $\left(P = \hat{\sigma}_0^2 \sum_{\overline{L}}^{-1}\right)$ مشاهدات متناسب با عکس ماتریس واریانس کوواریانس مشاهدات انتخاب شده است. فرموله کردن شرط می نیمم در اینجا نمی تواند بطور

مستقیم (آنطوریکه برای سرشکن پارامتریک انجام گرفت) انجام پذیرد. چون نمی توان بردار V را بطور صریح برحسب ماتریس B و بردار W (معادله 103-6) نوشت.

* اگر مدل غیر خطی باشد ، $F(L)=0$ ، از بسط آن به سری تیلور استفاده می کنیم :

$$F(L)=F\left(L^0\right)+\left.\frac{\partial F}{\partial L}\right|_{L=L^0}\left(L-L^0\right)+\ldots$$

در بسط فوق ، از ترمهای با درجات غیر خطی صرفنظر می کنیم. با قرار دادن $V=L-L^0$ و $B=\frac{\partial F}{\partial L}$ و $W=F\left(L^0\right)$ ، رابطه خطی $BV+W=0$ ایجاد می شود که همان فرمول 103-6 می باشد.

اما مسئله را می توان با کمک بردار k از r مؤلفه مجهول به نام ضرایب لاگرانژ حل کرد. می توان رابطه زیر را نوشت :

$$\min_{V\varepsilon R^n} V^T PV=\min_{V\varepsilon R^n}\left[V^T PV+2k^T(BV+W)\right] \qquad (6-105)$$

برای اینکه جمله داخل پرانتز طرف راست رابطه فوق ، مساوی صفر می باشد ، تابع فوق را بصورت زیر نشان می دهیم :

$$\Phi=V^T PV+2k^T(BV+W)$$

برای می نیمم کردن تابع فوق ، مشتق آنرا نسبت به بردار V مساوی صفر قرار می دهیم :

$$\frac{\partial\Phi}{\partial V}=2\hat{V}^T P+2k^T B=0$$

معادله فوق ، بعد از ترانسپوز کردن ، بصورت زیر در می آید :

$$P\hat{V}+B^T k=0$$

و از آن بردار $\hat{V}$ بصورت زیر محاسبه می شود :

$$\boxed{\hat{V}=-P^{-1}B^T k} \qquad (6-106)$$

سیستم معادلات فوق را معادلات لاگرانژ می نامند.

با جایگزینی معادله 106-6 در معادله 103-6 بردار V حذف می شود :

$$B\left(-P^{-1}B^T k\right)+W=0$$

و یا

$$\boxed{\left(BP^{-1}B^T\right)k=W} \qquad (6-107)$$

معادلات فوق را معادلات نرمال می نامند که معمولاً بصورت خلاصه زیر نوشته می شود :

$$Mk = W \tag{6-108}$$

که در آن

$$M = BP^{-1}B^T \tag{6-109}$$

می باشد. جواب معادلات نرمال 6-107 برای بردار k بصورت زیر می باشد :

$$k = M^{-1}W = \left(BP^{-1}B^T\right)^{-1}W \tag{6-110}$$

بعد از آنکه ضرایب K محاسبه شدند ، می توان مقدار باقیمانده ها را $(\hat{V})$ از فرمول 6-106 محاسبه کرد. سپس مشاهدات سرشکن شده $\hat{L}$ بصورت

$$\hat{L} = \bar{L} + \hat{V} \tag{6-111}$$

محاسبه می شوند.

در واقع ، اگر مورد نظر ما فقط بدست آوردن مشاهدات سرشکن شده باشد ، می توانیم از معادلات زیر که مقدار $\hat{L}$ را برحسب ماتریس های B و P و بردارهای $\bar{L}$ و C محاسبه میکنند استفاده کنیم :

$$\hat{L} = \bar{L} + \hat{V} = \bar{L} - P^{-1}B^T k = \bar{L} - P^{-1}B^T\left(BP^{-1}B^T\right)^{-1}\left(B\bar{L} - C\right) \tag{6-112}$$

و یا بصورت خلاصه زیر

$$\boxed{\hat{L} = (I - T)\bar{L} + HC} \tag{6-113}$$

که در آن I عبارتست از یک ماتریس یکه و

$$T = P^{-1}B^T\left(BP^{-1}B^T\right)^{-1}B$$

$$H = P^{-1}B^T\left(BP^{-1}B^T\right)^{-1} \tag{6-114}$$

مثال 6-20

مطلوبست حل مثال 6-16 با استفاده از روش سرشکنی معادلات شرط.

ما فقط یک معادله شرط ، بین مشاهدات (اختلاف ارتفاعات) $\bar{h}_i , i = 1,2,3$ داریم. همینطور میدانیم که تعداد درجات آزادی در آن مثال (16-6) برابر واحد می باشد. با قرار دادن $\Delta H = H_J - H_G$ ، معادله شرط ، بصورت زیر نوشته می شود :

$$\sum \bar{h}_i = \Delta H$$

که بفرم زیر تغییر داده می شود :

$$V_1 + V_2 + V_3 + \left(\bar{h}_1 + \bar{h}_2 + \bar{h}_3 - \Delta H\right) = 0$$

و یا بصورت ماتریسی

$$\underset{1\times3}{B}\ \underset{3\times1}{V} + \underset{1\times1}{W} = 0$$

که در آن

$$B = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \qquad \text{و} \qquad V = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} \qquad \text{و} \qquad W = \bar{h}_1 + \bar{h}_2 + \bar{h}_3 - \Delta H$$

می باشد. ماتریس وزن مشاهدات ، طبق مثال 16-6 بصورت

$$\underset{3\times3}{P} = diag\left(\frac{1}{d_1}, \frac{1}{d_2}, \frac{1}{d_3}\right)$$

داده می شود.

$$P^{-1} = diag\left(d_1, d_2, d_3\right)$$

سیستم معادلات نرمال برای محاسبه ضرایب k (معادلات 108-6) بصورت

$$\underset{1\times1}{M}\ \underset{1\times1}{k} = \underset{1\times1}{W}$$

که در آن

$$M = BP^{-1}B = d_1 + d_2 + d_3 = \sum_i d_i$$

خواهد بود. و از آنجا

$$k = M^{-1}W = \frac{1}{\sum_i d_i}\left(\sum_i \bar{h}_i - \Delta H\right)$$

باقیمانده های برآورد شده از معادله 106-6 محاسبه می شوند.

$$\hat{V} = -P^{-1}B^T k = -\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\left(\frac{\sum_i \bar{h}_i - \Delta H}{\sum_i d_i}\right) = -\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}\frac{\sum_i \bar{h}_i - \Delta H}{\sum_i d_i}$$

و یا

$$V_i = \frac{\Delta H - \sum \bar{h}_i}{\sum d_i} d_i \qquad و \qquad i = 1,2,3$$

نتایج فوق ، همانست که در مثال 16-6 با استفاده از روش سرشکنی پارامتریک بدست آمده بود. با توجه به اینکه $\Delta H = H_J - H_G$ می باشد ، در این مسئله بخصوص می بینیم که عمل سرشکنی باعث می شود که اختلاف $\left(\Delta H - \sum \bar{h}_i\right)$ به نسبت طول هر خط ترازیابی بین اختلاف ارتفاعهای مشاهده شده مربوطه تقسیم گردد ، یعنی به نسبت عکس MSE ها. مشاهدات سرشکن شده از فرمول 111-6 محاسبه می شوند :

$$\hat{L} = \bar{L} + \hat{V}$$

$$\hat{h}_i = \bar{h}_i + \hat{V}_i \qquad و \qquad i = 1,2,3$$

$$\hat{h}_i = \bar{h}_i + \frac{\Delta H - \sum \bar{h}_i}{\sum d_i} d_i$$

در خاتمه پارامترهای مجهول $\hat{x} = \left(\hat{H}_1, \hat{H}_2\right)$ را بوسیله ارتفاعات معلوم H_G, H_J و مشاهدات سرشکن شده $\hat{h}_i$ بصورت زیر محاسبه می کنیم :

$$\hat{H}_1 = H_G + \hat{h}_1 = H_G + \bar{h}_1 + \frac{d_1}{\sum d_i}\left(\Delta H - \sum \bar{h}_i\right)$$

$$\hat{H}_2 = H_J - \hat{h}_3 = H_J - \bar{h}_3 - \frac{d_3}{\sum d_i}\left(\Delta H - \sum \bar{h}_i\right)$$

نتایج بدست آمده فوق نظیرمقادیری است که در مثال 16-6 از طریق سرشکنی پارامتریک ، برای مجهولات H_2, H_1، بدست آمده بود.

مثال 21-6

مطلوبست حل مثال 17-6 با استفاده از روش سرشکنی معادلات شرط.

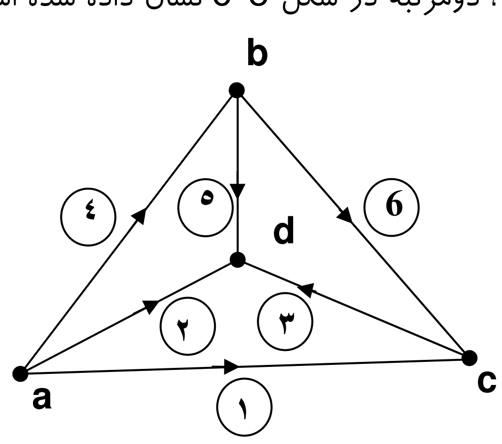
شکل شبکه ترازیابی مورد بحث در مثال 17-6 ، دومرتبه در شکل 8-6 نشان داده شده است.

در مثال فوق داریم :

$n = 6$ تعداد مشاهدات

$u = 3$ تعداد پارامترهای مجهول

که تعداد درجات آزادی $df = n - u = 3$ خواهد بود. خواهیم دید که ما می توانیم فقط 3 معادله مستقل شرط (به تعداد درجات آزادی) بین مشاهدات برقرار کنیم.



(شکل 8-6)

در شکل فوق چهار حلقه بسته (a-c-d-a) ، (a-d-b-a) ،
(b-c-d-b) و (a-c-b-a) وجود دارند ، یعنی می توان چهار معادله شرط ، بین مشاهدات نوشت. ولی سه تای آنها مستقل از همدیگر بوده و آن یکی نتیجه ای از آن سه معادله مستقل. بطور مثال حلقه چهارم از جمع سه حلقه اول بدست می آید. بطور مثال سه حلقه زیر را در نظر می گیریم :

$$I \text{ حلقه } = a-c-b-a$$

$$II \text{ حلقه } = a-c-d-a$$

$$III \text{ حلقه } = a-d-b-a$$

سه معادله شرط مربوط به سه حلقه فوق بصورت زیر نوشته می شوند :

$$\begin{aligned}
&\left(\bar{h}_1+V_1\right)-\left(\bar{h}_6+V_6\right)-\left(\bar{h}_4+V_4\right)=0\\
&\left(\bar{h}_1+V_1\right)+\left(\bar{h}_3+V_3\right)-\left(\bar{h}_2+V_2\right)=0\\
&\left(\bar{h}_2+V_2\right)-\left(\bar{h}_5+V_5\right)-\left(\bar{h}_4+V_4\right)=0
\end{aligned}$$

و یا

$$\begin{aligned}
&V_1-V_4-V_6+\left(\bar{h}_1-\bar{h}_4-\bar{h}_6\right)=0\\
&V_1-V_2+V_3+\left(\bar{h}_1-\bar{h}_2+\bar{h}_3\right)=0\\
&V_2-V_4-V_5+\left(\bar{h}_2-\bar{h}_4-\bar{h}_5\right)=0
\end{aligned}$$

معادلات شرط فوق دارای فرم ماتریسی

$$\underset{3\times6}{B}\ \underset{6\times1}{V}+\underset{3\times1}{W}=0$$

هستند که در آن

$$\underset{3\times6}{B}=\begin{bmatrix}1&0&0&-1&0&-1\\1&-1&1&0&0&0\\0&1&0&-1&-1&0\end{bmatrix} \quad \text{و} \quad \underset{1\times6}{V^T}=\begin{bmatrix}V_1&V_2&V_3&V_4&V_5&V_6\end{bmatrix}$$

$$\underset{3\times1}{W}=\begin{bmatrix}\left(\bar{h}_1-\bar{h}_4-\bar{h}_6\right)\\\left(\bar{h}_1-\bar{h}_2+\bar{h}_3\right)\\\left(\bar{h}_2-\bar{h}_4-\bar{h}_5\right)\end{bmatrix}$$

با قرار دادن مقادیر مشاهده شده ، از مثال 6-17 ، بردار فوق بصورت زیر در می آید :

$$\underset{3\times1}{W}=\begin{bmatrix}0\\0\\-0.1\end{bmatrix}\text{متر}$$

ماتریس وزن P ، با توجه به مثال 6-17 ، بصورت

$$\underset{6\times6}{P} = diag(0..25,0.5,0.5,0.25,0.5,0.25)$$

و

$$\underset{6\times6}{P^{-1}} = diag(4,2,2,4,2,4)$$

خواهد بود. معادلات نرمال برای ضرایب k بصورت زیر نوشته می شود :

$$\underset{3\times3}{M}\ \underset{3\times1}{k} = \underset{3\times1}{W}$$

که در آن

$$\underset{3\times3}{M} = BP^{-1}B^T = \begin{bmatrix} 12 & 4 & 4 \\ 4 & 8 & -2 \\ 4 & -2 & 8 \end{bmatrix}$$

بوده و معکوس آن بشکل

$$\underset{3\times3}{M^{-1}} = \begin{bmatrix} 0.15 & -0.1 & -0.1 \\ -0.1 & 0.2 & 0.1 \\ -0.1 & 0.1 & 0.2 \end{bmatrix}$$

خواهد بود و از آنجا

$$\underset{3\times1}{k} = M^{-1}W = \begin{bmatrix} 0.15 & -0.1 & -0.1 \\ -0.1 & 0.2 & 0.1 \\ -0.1 & 0.1 & 0.2 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 0.01 \\ -0.01 \\ -0.02 \end{bmatrix}$$

باقیمانده های برآورد شده عبارتند از :

$$\hat{V} = -P^{-1}B^T k = \begin{bmatrix} 0.00 \\ 0.02 \\ 0.02 \\ -0.04 \\ -0.04 \\ 0.04 \end{bmatrix} \text{متر}$$

مقادیر بدست آمده فوق برای باقیمانده ها همان است که از مثال 17-6 بدست آمده بود. و از آنجا مشاهدات سرشکن شده عبارت خواهند بود از :

$$\hat{L} = \overline{L} + \hat{V}$$

$$\begin{bmatrix}\hat{h}_1\\ \hat{h}_2\\ \hat{h}_3\\ \hat{h}_4\\ \hat{h}_5\\ \hat{h}_6\end{bmatrix}=\begin{bmatrix}6.16\\12.57\\6.41\\1.09\\11.58\\5.07\end{bmatrix}+\begin{bmatrix}0.00\\0.02\\0.02\\-0.04\\-0.04\\0.04\end{bmatrix}=\begin{bmatrix}6.16\\12.59\\6.43\\1.05\\11.54\\5.11\end{bmatrix}\text{متر}$$

پارامترهای مجهول (ارتفاع نقاط b و c و d) ، یعنی $\hat{H}_d, \hat{H}_c, \hat{H}_b$ بوسیله ارتفاع نقطه a ، H_a ، و مشاهدات سرشکن شده محاسبه می شوند :

$$\hat{H}_b = H_a + \hat{h}_4 = 0.0 + 1.05 = 1.05\text{متر}$$

$$\hat{H}_c = H_a + \hat{h}_1 = 0.0 + 6.16 = 6.16\text{متر}$$

$$\hat{H}_d = H_a + \hat{h}_2 = 0.0 + 12.59 = 12.59\text{متر}$$

نتایج فوق همانست که از طریق سرشکنی پارامتریک برای مجهولات فوق ، بدست آمده بود.
توجه کنید که در محاسبه ارتفاعات مجهول از ارتفاع نقطه a شما هر طریقی یا هر مسیری را در شکل مربوطه انتخاب کنید به یک نتیجه واحد خواهید رسید.

6-4-11 ماتریس واریانس کوواریانس نتایج در سرشکنی معادلات شرط

ماتریس واریانس کوواریانس مشاهدات سرشکن شده $\left(\sum \hat{L}\right)$ ، نتایج سرشکنی بطریقه کمترین مربعات معادلات شرط ، را می توان با اعمال قانون پخش کوواریانس(معادله 6-15) روی معادله 6-113 بدست آورد. در این معادله ماتریس های I و H و T ثابت هستند. همینطور C یک بردار ثابت می باشد و درنتیجه $\sum C$ ماتریس صفر خواهد بود. بنابراین می توان نوشت:

$$\sum \hat{L} = \left(\frac{\partial \hat{L}}{\partial \bar{L}}\right)\sum \bar{L}\left(\frac{\partial \hat{L}}{\partial \bar{L}}\right)^T = (I-T)\sum \bar{L}(I-T)^T$$

$$= \sum \bar{L} - \sum \bar{L}T^T - T\sum \bar{L} + T\sum \bar{L}T^T \qquad (6\text{-}115)$$

دو ماتریس $\left(\sum \bar{L}T^T\right)$ و $\left(T\sum \bar{L}\right)$ ماتریس های مربعی و قرینه می باشند. یکی از آنها ترانسپوز دیگری است. بنابراین با هم برابرند.

$$\sum \hat{L} = \sum \bar{L} - 2T\sum \bar{L} + T\sum \bar{L}T^T = \sum \bar{L} - T\sum \bar{L}\left(2I - T^T\right) \qquad (6\text{-}116)$$

با جایگزینی مقدار ماتریس T از فرمول 6-114 و دانستن اینکه $P = \sigma_0^2 \sum_{\bar{L}}^{-1}$ ، فرمول 6-116 بصورت زیر در می آید :

$$\sum \widehat{L} = \sigma_0^2 P^{-1} - \sigma_0^2 P^{-1}\left[B^T\left(BP^{-1}B^T\right)^{-1}B\right]P^{-1}\left\{2I - P^{-1}\left[B^T\left(BP^{-1}B^T\right)^{-1}B\right]^T\right\}$$

$$= \sigma_0^2 P^{-1}\left\{I - 2\left[B^T\left(BP^{-1}B^T\right)^{-1}B\right]P^{-1} + B^T\left(BP^{-1}B^T\right)^{-1}\left(BP^{-1}B^T\right)\left(BP^{-1}B^T\right)^{-1}BP^{-1}\right\}$$

(6-117)

با دانستن اینکه $\left(\left(BP^{-1}B^T\right)^{-1}\right)^4 = \left(BP^{-1}B^T\right)^{-1}$ ، فرمول فوق بصورت زیر در می آید:

(6-118)

$$\boxed{\sum \widehat{L} = \sigma_0^2 P^{-1}\left(I - B^T\left(BP^{-1}B^T\right)^{-1}BP^{-1}\right)}$$

در اینجا نیز نظیر سرشکنی پارامتریک برای بدست آوردن ماتریس واریانس کوواریانس برآورد شده از فاکتور واریانس برآورد شده $\hat{\sigma}_0^2$ بجای σ_0^2 استفاده می کنیم :

$$\hat{\sigma}_0^2 = \frac{\widehat{V}^T P \widehat{V}}{df} = \frac{\widehat{V}^T P \widehat{V}}{r} \qquad (6\text{-}119)$$

$$\hat{\sum} \widehat{L} = \hat{\sigma}_0^2 P^{-1}\left(I - B^T\left(BP^{-1}B^T\right)^{-1}BP^{-1}\right) \qquad (6\text{-}120)$$

و یا بصورت خلاصه تر

$$\boxed{\hat{\sum} \widehat{L} = \hat{\sigma}_0^2 (I - T) P^{-1}} \qquad * \qquad (6\text{-}121)$$

* به این طریق می توان نشان داد که :

$$\hat{\sum} \widehat{L} = \hat{\sigma}_0^2 T P^{-1}$$

نظیر سرشکنی پارامتریک ، می توان نشان داد که مقدار برآورد شده $\widehat{L}$ دارای ماتریس واریانس کوواریانس $\sum \widehat{L}$ می باشد که مجموع عناصر روی قطر اصلی آن می نیمم است. با در نظر گرفتن فرضیاتی که در بخش 6-4-8 عنوان شد می توان نشان داد که $\widehat{L}$ برآورد با ماکزیمم احتمال از L می باشد.

با در نظر گرفتن همبستگی مشاهدات سرشکن شده $\hat{L}$ ، ماتریس غیر قطری $\sum_{\hat{L}}$ ، میتوان شرایطی را که منجر به مشاهدات سرشکن شده مستقل $\hat{L}$ خواهد شد ، بصورت زیر بیان کرد :
1- اگر مشاهدات L مستقل باشند و 2- ماتریس B یک ماتریس متعامد باشد.
اگر شرایط فوق برقرار باشند ، ماتریسهای P^{-1} و T قطری خواهند بود. البته ممکن است سایر حالات ماتریس های فوق نیز منجر به مشاهدات مستقل $\hat{L}$ گردند. در خاتمه دوباره میتوان دید که ماتریس برآورد شده $\hat{\sum}_{\hat{L}}$ تحت تأثیر فاکتور انتخابی σ_0^2 نیست.

مثال 22-6

مطلوبست تعیین ماتریس واریانس کوواریانس برآورد شده $\hat{\sum}_{\hat{L}}$ برای سرشکنی معادلات شرط، در مثال 20-6 .

$$\hat{V}_i = \frac{\Delta H - \sum_i \bar{h}_i}{\sum_i d_i} d_i \qquad \text{و} \qquad i = 1,2,3$$

$$\underset{3\times3}{P} = diag\left(\frac{1}{d_1}, \frac{1}{d_2}, \frac{1}{d_3}\right)$$

از معلومات فوق که در مثال 20-6 بدست آمده بودند می توان محاسبات زیر را انجام داد :

$$\hat{\sigma}_0^2 = \frac{\hat{V}^T P \hat{V}}{r} = \frac{\left(\Delta H - \sum_i \bar{h}_i\right)^2}{\sum_i d_i}$$

$$\hat{\sum}_{\hat{L}} = \hat{\sigma}_0^2 (I - T) P^{-1}$$

برای بدست آوردن ماتریس فوق ، اول ماتریس $T = P^{-1} B^T M^{-1} B$ را محاسبه می کنیم. از مثال 20-6 داریم :

$$M^{-1} = \frac{1}{\sum_i d_i} \qquad \text{و} \qquad B = [1 \quad 1 \quad 1] \qquad \text{و} \qquad P^{-1} = diag(d_1, d_2, d_3)$$

$$T=\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\left[\frac{1}{\sum_i d_i}\right]\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \Rightarrow T=\frac{1}{\sum_i d_i}\begin{bmatrix} d_1 & d_1 & d_1 \\ d_2 & d_2 & d_2 \\ d_3 & d_3 & d_3 \end{bmatrix}$$

$$(I-T)P^{-1}=\begin{bmatrix} \alpha_1 & \alpha_1 & \alpha_1 \\ \alpha_2 & \alpha_2 & \alpha_2 \\ \alpha_3 & \alpha_3 & \alpha_3 \end{bmatrix}$$

که در آن $\alpha_j = d_j - \frac{d_j^2}{\sum_j d_j}, j=1,2,3$ می باشد. حال ماتریس $\hat{\sum}_{\hat{L}}$ بصورت زیر درمی آید :

$$\hat{\sum}_{\hat{L}}=\hat{\sigma}_0^2(I-T)P^{-1}=\left(\Delta H-\sum_i \bar{h}_i\right)\begin{bmatrix} \alpha_1 & \alpha_1 & \alpha_1 \\ \alpha_2 & \alpha_2 & \alpha_2 \\ \alpha_3 & \alpha_3 & \alpha_3 \end{bmatrix}$$

مثال 23-6

مطلوبست تعیین ماتریس $\hat{\sum}_{\hat{L}}$ برای شبکه ترازیابی مثال 21-6 که با روش معادلات شرط ، سرشکن شده است.

از مثال 21-6 داریم :

$$\underset{1\times 6}{\hat{V}^T}=[0.00,0.02,0.02,-0.04,-0.04,0.04]$$

$$\underset{1\times 6}{P}=diag(0.25,0.5,0.5,0.25,0.5,0.25)$$

$$r=df=n-u=6-3=3$$

$\hat{V}^T P\hat{V}=0.002$ مترمربع و $\hat{\sigma}_0^2=\frac{\hat{V}^T P\hat{V}}{r}=\frac{0.002}{3}=0.00067$ مترمربع

ماتریس مورد نظر $\hat{\sum}_{\hat{L}}$ از فرمول 121-6 محاسبه می شود :

$$\hat{\sum}_{\hat{L}}=\hat{\sigma}_0^2(I-T)P^{-1}$$

که در آن

$P^{-1}=diag(4,2,2,4,2,4)$ و $\underset{6\times 6}{T}=P^{-1}B^T M^{-1}B$

از مثال 21-6 داریم :

$$M^{-1}=\left(BP^{-1}B^{T}\right)^{-1}=\begin{bmatrix}0.15 & -0.1 & -0.1\\ -0.1 & 0.2 & 0.1\\ -0.1 & 0.1 & 0.2\end{bmatrix} \quad \text{و} \quad \underset{3\times 6}{B}=\begin{bmatrix}1 & 0 & 0 & -1 & 0 & -1\\ 1 & -1 & 1 & 0 & 0 & 0\\ 0 & 1 & 0 & -1 & -1 & 0\end{bmatrix}$$

$$B^{T}M^{-1}B=\begin{bmatrix}0.15 & -0.1 & 0.1 & -0.05 & 0 & -0.05\\ -0.1 & 0.2 & -0.1 & -0.1 & -0.1 & 0\\ 0.1 & -0.1 & 0.2 & 0 & -0.1 & 0.1\\ -0.05 & -0.1 & 0 & 0.15 & 0.1 & 0.05\\ 0 & -0.1 & -0.1 & 0.1 & 0.2 & -0.1\\ -0.05 & 0 & 0.1 & 0.05 & -0.1 & 0.15\end{bmatrix}$$

$$T=P^{-1}B^{T}M^{-1}B=\begin{bmatrix}0.6 & -0.4 & 0.4 & -0.2 & 0 & -0.2\\ -0.2 & 0.4 & -0.2 & -0.2 & -0.2 & 0\\ 0.2 & -0.2 & 0.4 & 0 & -0.2 & 0.2\\ -0.2 & -0.4 & 0 & 0.6 & 0.4 & 0.2\\ 0 & -0.2 & -0.2 & 0.2 & 0.4 & -0.2\\ -0.2 & 0 & 0.4 & 0.2 & -0.4 & 0.6\end{bmatrix}$$

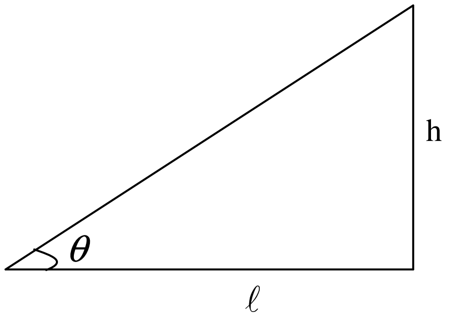
$$(I-T)P^{-1}=\begin{bmatrix}1.6 & -0.8 & 0.8 & -0.8 & 0 & -0.8\\ -0.8 & 1.2 & -0.4 & -0.8 & -0.4 & 0\\ 0.8 & -0.4 & 1.2 & 0 & -0.4 & 0.8\\ -0.8 & -0.8 & 0 & 1.6 & 0.8 & 0.8\\ 0 & -0.4 & -0.4 & 0.8 & 1.2 & -0.8\\ -0.8 & 0 & 0.8 & 0.8 & -0.8 & 2.4\end{bmatrix}$$

$$\hat{\sum}_{\hat{L}}=\hat{\sigma}_0^2(I-T)P^{-1}=10^{-4}\begin{bmatrix}10.72 & -5.36 & 5.36 & -5.36 & 0 & -5.36\\ -5.36 & 8.04 & -2.68 & -5.36 & -2.68 & 0\\ 5.36 & -2.68 & 8.04 & 0 & -2.68 & 5.36\\ -5.36 & -5.36 & 0 & 10.72 & 5.36 & 5.36\\ 0 & -2.68 & -2.68 & 5.36 & 8.04 & -5.36\\ -5.36 & 0 & 5.36 & 5.36 & -5.36 & 16.08\end{bmatrix}\text{مترمربع}$$

اگر ضریب 10^{-4} را از کنار ماتریس فوق برداریم ، اعداد داخل ماتریس به سانتیمتر مربع خواهند بود.مقایسه فاکتورواریانس برآورد شده $\hat{\sigma}_0^2$، فاکتورواریانس اولیه σ_0^2 وتوصیه هائی که در بخش 9-4-6 داده شد ، در اینجا نیز به قوت خویش باقیست.

5-6 تمرین شماره 6

1- برای تعیین ارتفاع یک دیوار (شکل زیر) ، فاصله افقی ℓ و زاویه قائم θ اندازه گیری شده اند و دارای میانگین

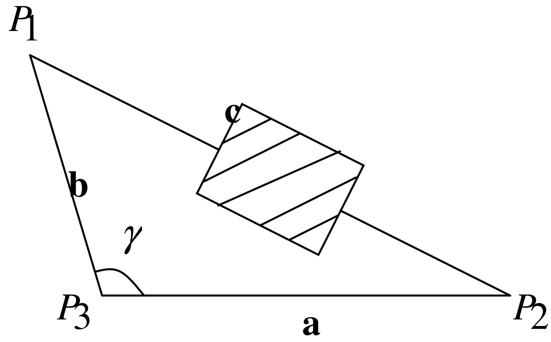


$\bar{\ell}=85.34$ متر $S_{\bar{P}}=2$ سانتیمتر

$\bar{\theta}=12^\circ 37' 30''$ $S_{\theta}=10''$

می باشند. مطلوبست تعیین ارتفاع h همراه با RMS آن.

2- برای تعیین فاصله $\overline{P_1P_2}=c$ که بطور مستقیم قابل اندازه گیری نیست ، بعلت وجود مانع، کمیتهای زیر اندازه گیری شده اند :



$\overline{P_2P_3}=\bar{a}=30$ متر $S_{\bar{a}}=3$ سانتیمتر

$\overline{P_1P_3}=\bar{b}=40$ متر $S_{\bar{b}}=4$ سانتیمتر

$\bar{\gamma}=60^\circ$ $S_{\bar{\gamma}}=25''$

مطلوبست تعیین فاصله $c=\overline{P_1P_2}$ همراه با انحراف معیار استاندارد آن تا نزدیکترین میلیمتر.

3- مطلوبست تعیین خطای استاندارد ارتفاع برج (h) مثال شماره 9 از تمرین شماره 4 (بخش 4-11) . فرض کنید که کمیتهای اندازه گیری شده β, α, ℓ و θ مستقل از همدیگرند.

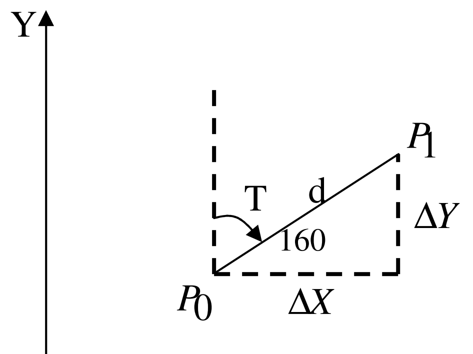
4- از نقطه P_0 در سیستم مختصات x, y شکل زیر فاصله $\bar{d}=5647.8$ متر و آزیموت $\bar{T}=49.9873$ گراد (صد گراد برابر نود درجه) به طرف نقطه P_1 اندازه گیری شده است. خطای نسبی طول $\bar{d}$ برابر 1.2×10^{-4} و خطای استاندارد آزیموت $\bar{T}$ برابر 0.08 سانتی گراد (یک گراد برابر صد سانتی گراد) می باشد. مطلوبست تعیین :

الف- اختلاف مختصات $\Delta\hat{Y}, \Delta\hat{X}$ بین نقاط P_1, P_0

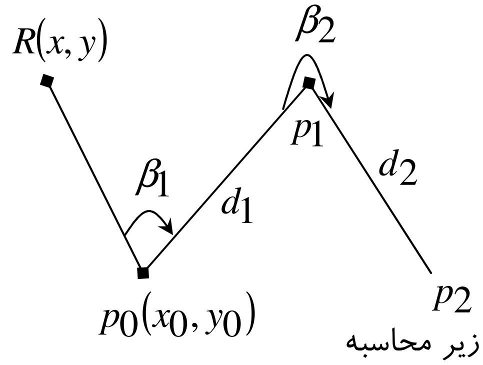
ب- ماتریس واریانس کوواریانس $\sum_{\hat{X}}$ برای $\hat{X}=\left(\Delta\hat{X}, \Delta\hat{Y}\right)$ به متر مربع

پ- RMS کمیتهای محاسبه شده $\Delta\hat{Y}, \Delta\hat{X}$ به ترتیب

ت- ضریب همبستگی بین $\Delta Y, \Delta X$



5- پیمایش زیر دارای دو ضلع $\overline{P_0P_1}$ و $\overline{P_1P_2}$ می باشد. مختصات (x_0, y_0) نقطه P_0 و مختصات (x, y) نقطه رفرانس R معلوم و بدون خطا یعنی مقادیر ثابت فرض شده اند. زوایای افقی β_2, β_1 همراه با فواصل افقی d_1 و d_2 اندازه گیری شده اند.



$x = 100$ متر $\qquad y = 200$ متر

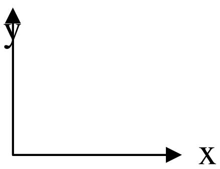
$x_0 = 150$ متر $\qquad y_0 = 150$ متر

$\bar{\beta}_1 = 75^\circ \qquad S_{\bar{\beta}_1} = 3''$

$\bar{\beta}_2 = 270^\circ \qquad S_{\bar{\beta}_2} = 2''$

$\bar{d}_1 = 100$ متر $\qquad \bar{d}_2 = 200$

خطای استاندارد فواصل اندازه گیری شده از فرمول زیر محاسبه می شوند :



$\delta_{\bar{d}}$ (سانتیمتر) $= 1.0$ (سانتیمتر) $+ d$ (متر) $\times 10^{-2}$

مطلوبست :

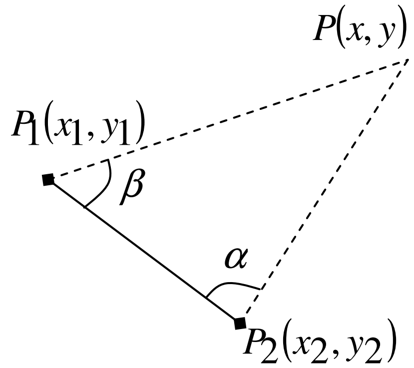
1- تعیین مختصات $(\hat{x}_1, \hat{y}_1)$ نقطه P_1 همراه با ماتریس واریانس کوواریانس $\sum(\hat{x}_1, \hat{y}_1)$

2- تعیین مختصات $(\hat{x}_2, \hat{y}_2)$ نقطه P_2 همراه با ماتریس واریانس کوواریانس $\sum(\hat{x}_2, \hat{y}_2)$

مختصات فوق را تا نزدیکترین میلیمتر و واریانس های مربوطه را تا نزدیکترین سانتیمتر مربع محاسبه کنید.

3- بحث روی همبستگی های موجود بین مختصات $\hat{y}_2, \hat{x}_2, \hat{y}_1, \hat{x}_1$

6- در یک مسئله تقاطع ، شکل زیر ، برای محاسبه مختصات (x, y) نقطه P از دو نقطه معلوم $P_1(x_1, y_1)$ و $P_2(x_2, y_2)$ زوایای افقی α و β اندازه گیری شده اند.



$\bar{x}_1 = 200$ متر

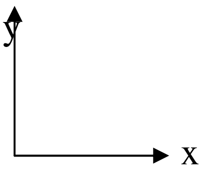
$\bar{y}_1 = 500$ متر $\qquad \sum(\bar{x}_1, \bar{y}_1) = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$ سانتیمترمربع

$\bar{x}_2 = 564/4$ متر

$$\bar{y}_2 = 300 \text{ متر} \qquad \sum(\bar{x}_2,\bar{y}_2) = \begin{bmatrix} 2 & -0.5 \\ -0.5 & 3 \end{bmatrix}$$

$$\bar{\alpha} = 90^\circ \qquad S_{\bar{\alpha}} = 3''$$

$$\bar{\beta} = 60^\circ \qquad S_{\bar{\beta}} = 2'' \qquad S_{\bar{\alpha}\bar{\beta}} = 0$$



مطلوبست تعیین مختصات $(\hat{x},\hat{y})$ نقطه P همراه با ماتریس واریانس کوواریانس $\sum(\hat{x},\hat{y})$ مربوطه.

7- فرض کنید که مشاهدات θ,ℓ در مسئله شماره 1 از این تمرین دارای خطاهای سیستماتیک 1- سانتی متر و 5 ثانیه می باشند. مطلوبست تعیین خطای کل روی ارتفاع محاسبه شده h به سانتیمتر.

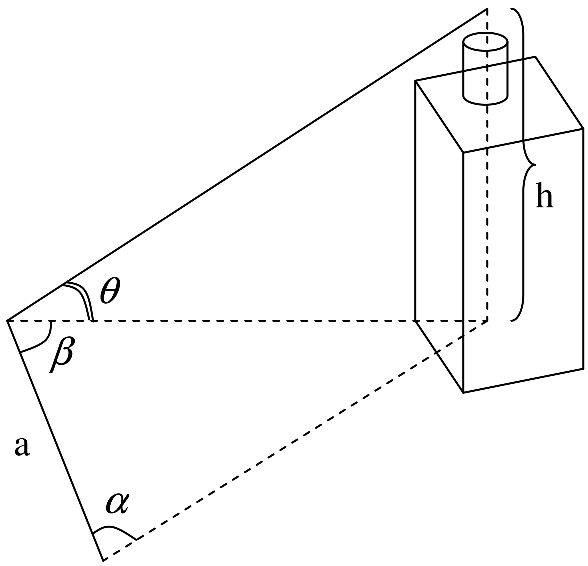
8- خطای موجود در جمع صد عدد ار در حالتهای زیر پیدا کنید :

الف- هر کدام از اعداد تا رقم سوم بعد از ممیز گرد شده است.

ب- ارقام اعشاری بعد از رقم سوم در هر کدام از اعداد حذف شده اند.

نتایج حاصل در دو حالت فوق را با هم مقایسه کنید.

9- برای تعیین ارتفاع (h) یک برج ، روش نشان داده شده در شکل زیر پیشنهاد شده است. زوایای θ,β,α و طول a اندازه گیری شده اند. مقادیر تقریبی کمیتهای اندازه گیری شده عبارتند از :



$\alpha = 60^\circ$ و $\beta = 30^\circ$ و $\theta = 45^\circ$ و

$a = 100$ متر

دقت اندازه گیری زوایای افقی β,α با هم مساوی و برابر $S_{\bar{\alpha}} = S_{\bar{\beta}} = 2''$ می باشند. با در نظر گرفتن اینکه خطاهای ناشی از طول a و زاویه θ روی ارتفاع h با هم برابر فرض

شده اند و به منظور رسیدن به خطائی مساوی یا کمتر از 2.5 سانتیمتر برای ارتفاع h ، دقت مورد لزوم در اندازه گیری طول a و زاویه θ را تعیین کنید.

10- در یک مثلث بندی مسطحاتی قرار است زوایای افقی با دو نوع مختلف تئودولیت I و II اندازه گیری شوند. برای امتحان تئودولیت ها زاویه ای را چندین بار بوسیله هر کدام از آنها اندازه گیری کرده ایم. خطای استاندارد یک مشاهده (خطای استاندارد نمونه) بر ای تئودولیت I برابر $S_{L_1} = 1.5''$ و برای تئودولیت II برابر $S_{L_2} = 2.5''$ بدست آمد. اگر خطای مجاز در اندازه گیری زوایای شبکه مثلث بندی 0.5 ثانیه باشد پیدا کنید تعداد تکرار اندازه گیری را : الف- اگر تئودولیت I بکار برده شود. ب- اگر تئودولیت II بکار گرفته شود.

11- طول یک میله آهنی چندید بار اندازه گیری و در زیر ثبت شده است :

3.321 و 3.310 و 3.273 و 3.253 و 3.302 و 3.284 و 3.304 و 3.295 و 3.263 و 3.270 متر

مطلوبست :

1- تعیین طول میله (میانگین)

2- خطای استاندارد یک مشاهده (RMS)

3- خطای استاندارد میانگین مشاهدات

12- طول ℓ را در پنج دور مختلف وهردور به دفعات اندازه گیری کرده و یانگین اندازه ها را همراه با خطای استاندارد مربوطه در جدول زیر نوشته ایم.

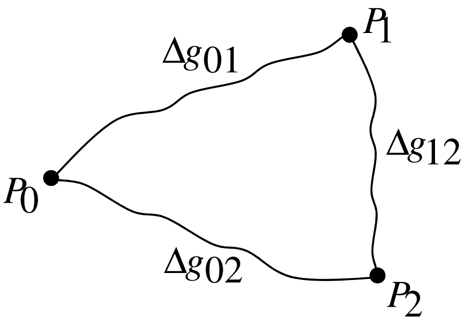
روز	دوشنبه	سه شنبه	چهارشنبه	پنجشنبه	جمعه
$\bar{\ell}_i$	101.01 متر	100.0	99.97	99.96	100.02
$S_{\bar{\ell}_i}$	2 سانتیمتر	1	4	5	3

مطلوبست تعیین طول ℓ (میانگین وزن دار) ، $\hat{\ell}$ ، همراه با خطای استاندارد آن یعنی $S_{\hat{\ell}}$.

13- شبکه گراویمتری (شتاب ثقل سنجی) زیر داده شده است. شتاب ثقل های g_2, g_1 را در نقاط P_2, P_1 همراه با ماتریس واریانس کوواریانس آنها پیدا کنید. مقدار شتاب ثقل در نقطه P_0

برابر $g_0 = 979832.12$ میلی گال (هزار میلی گال برابر یک گال و برابر یک سانتیمتر از مجذور ثانیه می باشد).

در جدول زیر اختلافهای شتابهای اندازه گیری شده همراه با مدت اندازه گیری هر اختلاف نوشته شده است.



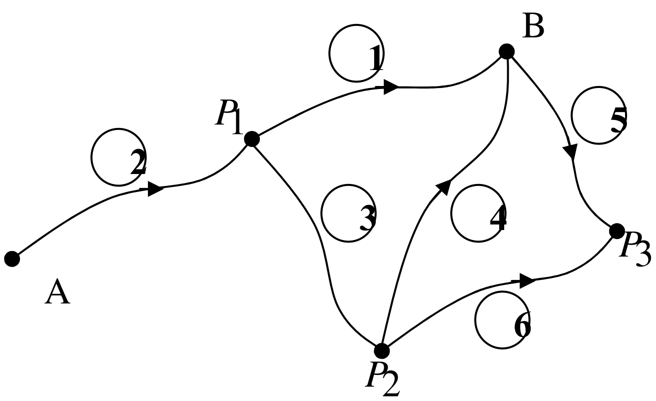
ایستگاه از	ایستگاه به	Δg (میلی گال)	ΔT (ساعت)
P_0	P_1	-9.82	2.5
P_2	P_0	-27.78	1.5
P_1	P_2	+38.42	2.0

فرض می کنیم که Δgها مستقل از همدیگر بوده و واریانس آنها متناسب با زمان اندازه گیری ΔT باشند.

14- شبکه ترازیابی زیر مفروض است. ارتفاع نقاط A و B معلوم و بدون خطا می باشند.

$H_A = 300.00$ متر و $H_B = 302.245$ متر

در جدول زیر اختلافات مشاهده شده $(\bar{h}_i)$ همراه با طول خط ترازیابی داده شده است.



شماره خط ترازیابی	ایستگاه از	ایستگاه به	$\bar{h}_i$ (متر)	ℓ_i (کیلومتر)
1	P_1	B	1.245	1.0

2	A	P_1	0.990	0.5
3	P_1	P_2	0.500	1.0
4	P_2	B	0.751	1.0
5	P_3	B	1.486	0.5
6	P_3	P_2	0.740	1.5

در شکل فوق ، فلش ها در جهت افزایش ارتفاعات تراز داده شده اند. مشاهدات فوق ، مستقل از هم فرض شده و بازای واریانس متناسب با طول خط ترازیابی می باشد. شبکه ترازیابی فوق را به روش پارامتریک سرشکن کرده و موارد زیر را جواب دهید :

1- ارتفاع نقاط P_1 و P_2 و P_3 ، یعنی $\hat{H}_1$ و $\hat{H}_2$ و $\hat{H}_3$ را بدست آورید.

2- اختلاف ارتفاعات (مشاهدات) سرشکن شده را بدست آورید.

3- فاکتور واریانس برآورد شده $\hat{S}_0^2$ را مجاسبه کرده و آنرا با فاکتور واریانس اولیه S_0^2 مقایسه کنید.

4- ماتریس واریانس کوواریانس $\hat{\sum}\hat{X}$ برآورد شده مجهولات $\hat{X}=\left(\hat{H}_1,\hat{H}_2,\hat{H}_3\right)$ را بدست آورید.

15- شبکه ترازیابی مسئله 14 را با روش معادلات شرط ، سرشکن کرده به سؤالات مندرج در آن مسئله جواب دهید و بجای سؤال چهارم ماتریس $\hat{\sum}\hat{L}$ را پیدا کنید. نتایج بدست آمده از هر دو مسئله را با هم مقایسه کنید.

16- دو کمیت فیزیکی Y و Z طبق مدل خطی زیر به هم مربوط می باشند.

$$Z=\alpha Y+\beta$$

که در آن α و β دو مقدار ثابت می باشند. برای تعیین دو ضریب α و β مشاهدات روی کمیتهای Y و Z انجام شده و در جدول زیر تنظیم شده است.

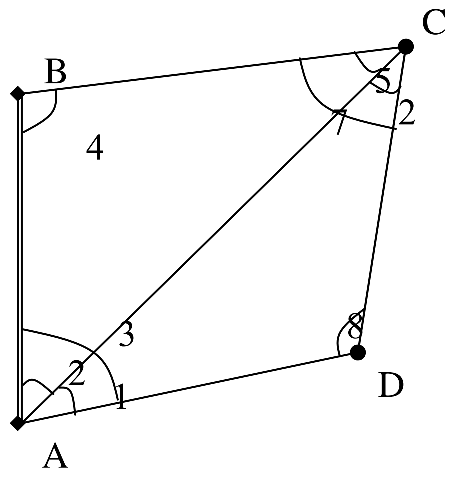
$\bar{Y}$	1	3	4	6	8	9	11	14
$\bar{Z}$	1	2	4	4	5	7	8	9

با فرض اینکه $\bar{Y}$ها بدون خطا و $\bar{Z}$ها با دقت یکسان اندازه گیری شده اند ، مطلوبست برآورد $\hat{\alpha}$ و $\hat{\beta}$ بطوریکه بهترین خط ، از نقطه نظر کمترین مربعات ، منطبق بر نقاط فوق را ایجاد کند.

17- مطلوبست حل مسئله شماره 16 ، این بار فرض بر اینست که $\bar{Z}$ها بدون خطا و $\bar{Y}$ها با دقت یکسان اندازه گیری شده اند. نتیج حاصله را ($\hat{\alpha}$و$\hat{\beta}$) با نتایج مسئله 16 مقایسه کنید.

18- شکل زیر یک شبکه مثلث بندی را با بار ثابت $AB=2$ کیلومتر نشان می دهد. هشت زاویه شماره گذاری شده اندازه گیری شده اند. هر کدام از زاویه ها به تعداد مختلف n_i ، مندرج در جدول زیر، تکرار شده اند.

شماره زاویه	n_i	میانگین زاویه
1	2	$82°7'9.5''$
2	2	$28°22'17.70''$
3	5	$110°29'25.02''$
4	3	$125°53'23.67''$
5	2	$25°44'9.30''$
6	2	$29°11'17.5''$
7	5	$55°3'29.32''$
8	3	$68°33'32.33''$



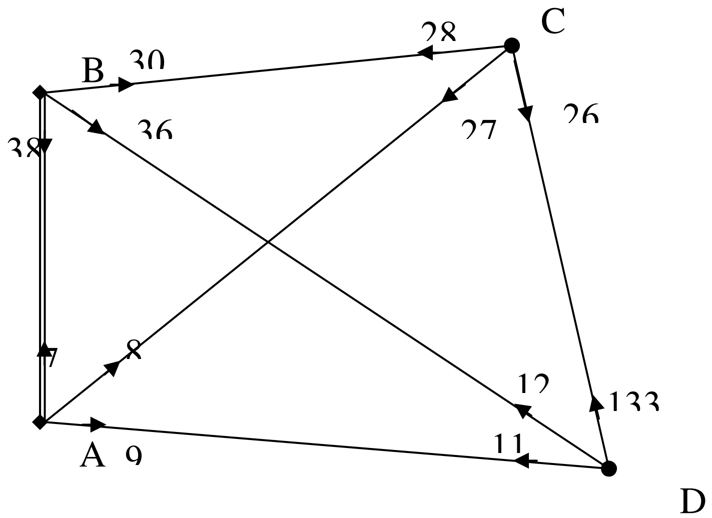
فرض می کنیم که تمام زوایا با یک دستگاه و در شرایط یکسان اندازه گیری شده اند. یعنی وزن هر زاویه متناسب با تعداد تکرار n_i می باشد.

مطلوبست :

1- تعیین طول $C\hat{}D$ با استفاده از مشاهدات سرشکن شده (زوایای سرشکن شده) وبدون در نظر گرفتن اضافه کرویت مثلث ها.

2- با در نظر گرفتن بار ثابت (بدون خطا) AB ، خطای نسبی طول CD را پیدا کنید.

19- شکل زیر یک شبکه چهار ضلعی با دو قطر را نشان می دهد. امتدادهائی که با فلش نشان داده شده اند با دقت یکسان قرائت شده اند. باز $AB=25$ کیلومتر و بدون خطا در نظر گرفته شده است. مقدار اضافه کرویت در هر مثلث بطور تقریب بصورت زیر داده شده است:



$\overset{\Delta}{ABC}$ $\varepsilon=3.126''$

$\overset{\Delta}{ABD}$ $\varepsilon=1.556''$

$\overset{\Delta}{CBD}$ $\varepsilon=3.085''$

$\overset{\Delta}{ACD}$ $\varepsilon=1.515''$

نتایج حاصل از قرائت امتدادها در جدول زیر خلاصه شده اند.

ایستگاه تئودولیت	ایستگاه نشانه روی	شماره امتداد	مقدار قرائت شده
A	B	7	$0^\circ 0' 0''$
	C	8	$91^\circ 30' 30.35''$
	D	9	$125^\circ 54' 33.91''$
B	C	30	$0^\circ 0' 0''$

	D	36	$28^\circ 22' 12.26''$
	A	38	$110^\circ 29' 27.13''$
C	D	26	$0^\circ 0' 0''$
	A	27	$29^\circ 19' 17.52''$
	B	28	$35^\circ 3' 26.80''$
D	A	11	$0^\circ 0' 0''$
	B	12	$35^\circ 7' 29.00''$
	C	13	$68^\circ 33' 32.60''$

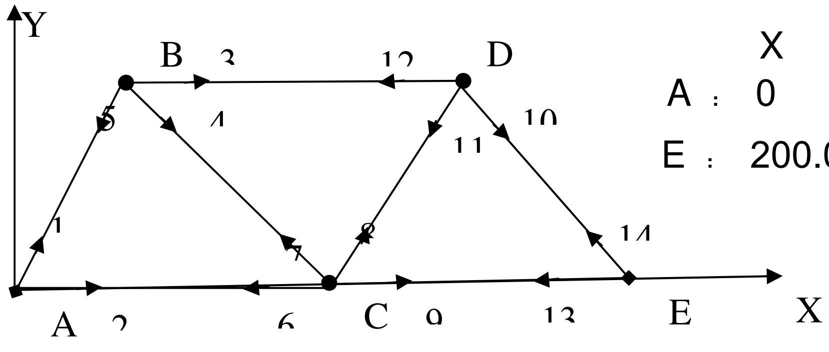
مطلوبست :

1- یک سرشکنی با معادلات شرط تشکیل داده ، مشاهدات سرشکن شده را همراه با ماتریس واریانس کوواریانس آنها $\left(\sum_{\hat{L}}\right)$ پیدا کنید.

2- با استفاده از تئوری لژاندر برای یک مثلث کروی یعنی با تفاضل یک سوم اضافه کرویت هر مثلث از زوایای آن مثلث و با در نظر گرفتن مثلث های جدید بعنوان مثلثهای مسطحه ، طول $C^\wedge D$ را با استفاده از باز AB و امتدادهای سرشکن شده بدست آورید و صحت آنرا با محاسبه آن از طریق دیگر کنترل کنید.

3- خطای نسبی برآورد شده طول CD را محاسبه کنید.

20- شکل زیر یک شبکه مثلث بندی را با دو نقطه معلوم و ثابت (بدون خطا) A و E که مختصات آنها در زیر داده شده است نشان می دهد.



	X		Y	
A :	0	متر	0	متر
E :	200.0	متر	0	متر

چهارده امتداد نشان داده شده (با فلش) با دقت یکسان قرائت شده اند و نتایج حاصل در جدول زیر آمده است.

ایستگاه قراول روی	ایستگاه نشانه روی	شماره امتداد	امتدادهای مشاهده شده
A	B	1	$0°0'0''$
	C	2	$60°0'10.0''$
B	D	3	$0°0'0''$
	C	4	$60°0'5.0''$
	A	5	$119°59'50.0''$
C	A	6	$0°0'0''$
	B	7	$59°59'55.0''$
	D	8	$120°0'0''$
	E	9	$180°0'51.0''$
D	E	10	$0°0'0''$
	C	11	$59°59'45.0''$
	B	12	$119°59'55.0''$
E	C	13	$0°0'0''$
	D	14	$60°0'15.0''$

مطلوبست تنظیم یک برنامه کامپیوتری که شبکه فوق را به روش پارامتریک (معادلات مشاهدات) سرشکن کرده و مختصات نقاط B و C و D را محاسبه نماید. در معادلات

مشاهدات مربوطه ، امتدادها بعنوان مشاهدات در نظر گرفته شوند ، نه زاویه ها. موارد زیر را به ترتیب اولویت در برنامه نویسی فوق در نظر بگیرید :

1- تعیین تعداد پارامترهای مجهول و تعداد درجات آزادی
2- مدل غیر خطی مشاهدات
3- تعیین تقریبی پارامترهای مجهول (مختصات نقاط B و C و D)
4- فرم خطی مدل ریاضی ، یعنی $V = A\Delta X - \Delta\bar{L}$ با انتخاب اسامی مناسب برای بردارهای V و ΔX و معرفی مقادیر عددی بردار $\Delta\bar{L}$ و تعیین مقادیر عددی عناصر ماتریس ضرایب A
5- تشکیل ماتریس واریانس کوواریانس مشاهدات $\left(\sum_{\Delta\bar{L}}\right)$ با فرض اینکه خطای استاندارد قرائت هر امتداد برابر 2 ثانیه باشد.

ضمیمه شماره I

<u>مفروضات بکار رفته در تابع PDF گوس و بدست آوردن تابع PDF گوس</u>

روش بدست آوردن تابع گوس که در زیر می آید منسوب به G.H.L.Hagen(1837) میباشد. قانون نرمال ، اولین بار توسط De Maives(1733) بصورت فرمول بیان شده است.

(1) فرض کنیم که تعداد m علت فیزیکی مستقل در اندازه گیری یک کمیت دخالت دارند و فرض کنیم که هر علت ، باعث یک خطای جزئی $+\Delta$ یا $-\Delta$ در اندازه کمیت که دارای خطای کل ε می باشد می شود. خطای کل ε می تواند بصورت یکی از ترکیب های (تعداد ترکیب ها $n = 2^m$ می باشد) m خطای جزئی $\pm\Delta$ بیان شود. از اول مشاهده می شود که خطای ε در محدوده $< -m\Delta, +m\Delta >$ بوده و در ضمن نمی تواند هر مقدار دلخواهی را در محدوده فوق داشته باشد ، بلکه فقط مقادیری را که از ترکیب m خطای $\pm\Delta$ بدست می آیند خواهد داشت. بسادگی دیده می شود که دو مقدار متوالی ε در محدوده فوق دارای اختلاف 2Δ میباشند ، برای اینکه یکی از آنها با جابجائی $-\Delta$ با $+\Delta$ و یا برعکس ، در دیگری بدست می آید. با تقسیم عرض (Range) ε ،

$$Ra(\varepsilon) = m\Delta - (-m\Delta) = 2m\Delta$$

بر نمو ε یعنی 2Δ معلوم می گردد که ε یکی از m+1 خطای زیر

$$\varepsilon_i = (2i - m)\Delta \qquad i = 0,1,\ldots,m \qquad (I-1)$$

را دارا می باشد. ε_iها حاصل ترکیبات ممکنه m خطای جزئی $\pm\Delta$ می باشند.

(2) فرض می کنیم که مجموعه D (مجموعه تمام مقادیر مجاز ε)

$$D \equiv \{\varepsilon_0, \varepsilon_1, \ldots \varepsilon_i, \ldots \varepsilon_m\}$$

فضای احتمال نمونه تصادفی ε که ، از تعداد 2^m ترکیب های مختلف m خطا تشکیل شده است ، باشد. بدیهی است که تعداد زیادی از 2^m ترکیب با هم برابرند. زیرا د رآنجا فقط تعداد m+1 تا ε وجود دارد که از نظر مقدار متفاوتند $(I-1)$. اگر شمارش یک مقدار بخصوص ε_i را با C_i نشان دهیم (بخش 1-1-3) مقدار آن از طریق احتمال مرکب بصورت زیر محاسبه می شود (بخش 3-2) :

$$C_i = \binom{m}{i} = \frac{m(m-1)(m-2)\ldots(m-i+1)}{i(i-1)(i-2)\ldots 1} = \frac{\prod_{j=m-i+1}^{m} j}{\prod_{j=1}^{i} j} \qquad (I-2)$$

و مقدار احتمال واقعی هر کدام از ε_iها بصورت

$$P(\varepsilon_i) = \frac{C_i}{n} = \frac{\binom{m}{i}}{2^m} \qquad (I-3)$$

خواهد بود(بخش 2-1-3) .

(3) فرمول فوق شکل تابع PDF واقعی نمونه ε را در فضای احتمال گسسته D مشخص می کند. از آنجائیکه هدف نهائی ما پیدا کردن یک فرمول ریاضی برای تابع PDF پیوسته مربوط به تابع PDF گسسته فوق می باشد (بعداً در این مورد شرح داده می شود) می باشد ، لازم است که تابع احتمال P ، علاوه بر آنکه برحسب i مشخص شده است $(I-3)$ ، برحسب ε_i نیز بیان گردد. برای اینکار ساده ترین راه ، استفاده از اختلافهای محدود می باشد. اختلاف $\delta P(\varepsilon_i)$ را بصورت زیر تعریف می کنیم :

$$\delta P(\varepsilon_i) = P(\varepsilon_i) - P(\varepsilon_{i-1})$$

از فرمول $(I-3)$ خواهیم داشت :

$$\delta P(\varepsilon_i) = \binom{m}{i}\Big/2^m - \binom{m}{i-1}\Big/2^m = \left[\binom{m}{i} - \binom{m}{i}\frac{i}{m-i+1}\right]\Big/2^m$$

$$= \frac{\binom{m}{i}}{2^m}\left(1 - \frac{i}{m - i + 1}\right)$$

بدیهی است که نسبت $\frac{\delta P(\varepsilon_i)}{P(\varepsilon_i)}$ بصورت زیر خواهد بود :

$$\frac{\delta P(\varepsilon_i)}{P(\varepsilon_i)} = 1 - \frac{i}{m - i + 1} \qquad (I-4)$$

از طرف دیگر i را می توان برحسب ε_i از فرمول $(I-1)$ بدست آورد.

$$\varepsilon_i = (2i - m)\Delta$$

$$i = \frac{1}{2}\left(\frac{\varepsilon_i}{\Delta} + m\right)$$

با در نظر داشتن $\delta P(\varepsilon) = 2\Delta$ و فرمول فوق ، معادله $(I-4)$ بصورت زیر در می آید :

$$\frac{\delta P(\varepsilon)}{P(\varepsilon)} = 1 - \frac{\frac{\varepsilon}{\delta\varepsilon} + \frac{m}{2}}{m - \frac{\varepsilon}{\delta\varepsilon} - \frac{m}{2} + 1} = \frac{1 + \frac{m}{2} - \frac{\varepsilon}{\delta\varepsilon} - \frac{\varepsilon}{\delta\varepsilon} - \frac{m}{2}}{1 + \frac{m}{2} - \frac{\varepsilon}{\delta\varepsilon}}$$

$$= \frac{1 - 2\frac{\varepsilon}{\delta\varepsilon}}{1 + \frac{m}{2} - \frac{\varepsilon}{\delta\varepsilon}} = \frac{2\varepsilon - \delta\varepsilon}{\left(1 + \frac{m}{2}\right)\delta\varepsilon - \varepsilon}$$

(4) قدم بعدی برگردانیدن تابع PDF گسسته به تابع PDF پیوسته مربوطه می باشد ، یعنی بدست آوردن تابع PDF پیوسته مربوطه. فرض می شود که متغیر ε چنان باشد که مثل متغیر گسسته ε در قسمت (1) تعریف شده است. با این تفاوت که m مجاز باشد تا بینهایت تغییر کند ، یعنی $m \to \infty$. اما با تغییر m تا بینهایت ، قدر مطلق ε نیز تا بینهایت تغییر خواهد کرد (معادله $(I-1)$) . ولی این برخلاف تجربه ماست که براساس آن خطاها دارای اندازه محدود می باشند. بنابراین ما باید فرض دیگری را قرار دهیم و آن عبارتست از اینکه هر چه تعداد m افزایش می یابد قدر مطلق خطاهای جزئی (Δ) بطرف صفر میل می کند ، بطوریکه حاصلضرب $m\Delta$ درمعادله $(I-1)$ همیشه محدود باقی می ماند.

با قبول دو فرض فوق ، معادله اختلاف محدود را بصورت

$$\lim_{\substack{m \to \infty \\ \delta\varepsilon \to 0}} \frac{\delta P(\varepsilon)}{P(\varepsilon)} = -\lim_{\substack{m \to \infty \\ \delta\varepsilon \to 0}} \frac{2\varepsilon - \delta\varepsilon}{\left(1 + \frac{m}{2}\right)\delta\varepsilon - \varepsilon} \qquad (I-5)$$

می نویسیم. معادله فوق یک معادله دیفرانسیل معمولی برای تابع $P(\varepsilon)$ پیوسته می باشد و بنابراین می توان نوشت :

$$\frac{dP\left(\varepsilon\right)}{P\left(\varepsilon\right)}=-\frac{2\varepsilon-d\varepsilon}{\left(1+\frac{m}{2}\right)d\varepsilon-\varepsilon}$$

برای سهولت در حل معادله فوق ، هر دو صورت و مخرج طرف راست معادله فوق را در $d\varepsilon$ ضرب کرده و فرض می کنیم که $md\varepsilon^2$ ثابت است و بعلاوه فرض می کنیم که

$$d\varepsilon^2 << \varepsilon d\varepsilon << md\varepsilon^2 = C$$

سپس می توان نوشت :

$$\frac{dP}{P}\cong-\frac{2\varepsilon d\varepsilon}{\frac{C}{2}}=-\frac{4}{C}\varepsilon d\varepsilon \qquad (I-6)$$

(5) سرانجام می توان معادله دیفرانسیل فوق را با انتگرال گیری مستقیم حل کرد.

$$\int\frac{dP}{P}\cong-\frac{4}{C}\int\varepsilon d\varepsilon+\text{مقدارثابت}$$

$$LnP\cong-\frac{4}{C}\frac{\varepsilon^2}{2}+\text{مقدارثابت}$$

مقدار ثابت را برابر LnK (لگاریتم طبیعی عدد ثابت K) قرار می دهیم. معادله فوق بشکل زیر در می آید :

$$P\cong Ke^{-\frac{2\varepsilon^2}{C}} \qquad (I-7)$$

حال این سؤال پیش می آید که آیا می توان هر دو ثابت C و K را بعنوان دو پارامتر مستقل تابع فوق در نظر گرفت. می دانیم که شرط اساسی برای یک تابع احتمال عبارتست از :

$$\int_{-\infty}^{+\infty}P\left(\varepsilon\right)d\varepsilon=1 \qquad (I-8)$$

با قرار دادن $(I-7)$ در شرط فوق ، خواهیم داشت :

$$\int_{-\infty}^{\infty}P\left(\varepsilon\right)d\varepsilon=\int_{-\infty}^{\infty}Ke^{-\frac{2\varepsilon^2}{C}}d\varepsilon=K\int_{-\infty}^{\infty}e^{-\frac{2\varepsilon^2}{C}}d\varepsilon=1$$

و از آنجا

$$K=\frac{1}{\int_{-\infty}^{\infty}e^{-\frac{2\varepsilon^2}{C}}d\varepsilon}$$

جواب سؤال فوق اینست که K نباید بعنوان یک پارامتر مستقل در نظر گرفته شود ، بلکه K تابعی است از پارامتر C (معادله فوق) و مقدار آن از حل انتگرال فوق بدست می آید.

$$\int_{-\infty}^{\infty} e^{-\frac{2\varepsilon^2}{C}} d\varepsilon = 2\int_{0}^{\infty} e^{-\frac{2\varepsilon^2}{C}} d\varepsilon = \sqrt{\frac{Cx}{2}} \qquad (I-9)$$

و از آنجا

$$K = \sqrt{\frac{2}{Cx}} \qquad (I-10)$$

در نتیجه تابع PDF گوس به شکل زیر نوشته می شود :

$$\boxed{P(\varepsilon) = G(C,\varepsilon) = \sqrt{\frac{2}{Cx}} e^{-\frac{2\varepsilon^2}{C}}} \qquad (I-11)$$