تخمین پارامترهای احتمال:

۱-۴ : روش احتمال شرطی

اجازه دهید(X1,Y1) , … , (Xn,Yn) نشان دهنده نمونه های تصادفی از جامعه n باشند این نمونه ها برای تخمین Pr(C|A) استفاده می شوند .

احتمال شرطی رخداد C به شرط رخداد A به وسیله فرمول اماری زیر محاسبه می شود :

$$p(C \mid A) = \frac{\sum_{i=}^{n} 1xA(x_i)xc(y_i)}{\sum_{i=}^{n} 1XA(x_i)} \quad (۴ \;\; .۱)$$

که وظایف مشخصه های Xc, XA نشان داده می شوند به وسیله :

$$x_A(x) = \begin{cases} 1 & if x \in A \\ 0 & otherwise \end{cases} \quad (۴ \;.۲)$$

$$x_C(y) = \begin{cases} 1 & if \quad y \in C \\ 0 & otherwise \end{cases} \quad (۴ \;.۳)$$

حالافرض کنید به جای پدیده های معمولی A و C پدیده های فازی جایگزین شوند .

این به این معناست که به وسیله **mfs** پدیده های **A,C** به **μA** و **μC** تعریف

شوندو

به جای **XA،Xc** در معادله ۴.۱ جایگزین شوند . در نتیجه خواهیم داشت :

$$\hat{p}(C \mid A) = \frac{\sum_{i=1}^{n} \mu_A(x_i)\mu c(y_i)}{\sum_{i=1}^{n} \mu_A(x_i)} \quad (۴.۴)$$

این فرمول پایه تعریف احتمال رخداد در پدیده فازی می باشد (درس ۳۷) .

مشتق اول فرمول ۴.۴ درسهای ۳۵و۳۶ را پدید می آورد .

نتیجه فرمول ۴.۴ در تخمین پارامترهای شرطی در**PFS** استفاده می شود .

این دیدگاه دردرسهای ۱۶و۱۸و۳۴ دنبال می شود که به روشهای احتمال

شرطی در این تز اشاره

می کند .

فرض کنید مجموعه اطلاعاتی شامل **n** نمونه به صورت ( ) **(i=1,2, ...,n)**

Xi,Yi

برای تخمین پارامترهای احتمال در دسترس باشد همچنین فرض کنید که

هم مقدمه وهم نتیجه **mfs** درسیستم تعیین شده است ونیاز به بهینه سازی

بیشتر نمی باشد یعنی فقط پارامترهای احتمال درتخمین باقی بمانند . به نظر

منطقی می آید که پارامترهای **Pj,k** واقعی رابرای تخمین احتمال شرطی

پدیده فازی **Ck** به شرط رخداد پدیده فازی **Aj** قرار دهیم . اگرچه ورودی **X**

به تعریف بیشتر احتیاج ندارد اما برای نشان دادن غیر عادی بودن محاسبات **mfμAj** و **mfμ¯Aj** باید ازفرمول زیراستفاده شود :

$$\hat{p}_{j,k} = \frac{\sum_{i=1}^{n} \overline{\mu}_A(x_i)\mu c_k(y_i)}{\sum_{i=1}^{n} \overline{\mu}_{A\,j}(x_i)} \qquad (۴.۵)$$

بنابراین **Pj,k** واقعی است و برای تخمین احتمال شرطی پدیده فازی **Ck** ونشان دادن غیر عادی بودن پدیده فازی **Aj** باید ازآن استفاده شود .

توجه داشته باشید که **PFSs** برای نمونه های برگشتی یک قانون پایه دارد که فقط با همان قانون که در پارامترهای شرطی **Pj,k** استفاده می شود ودرفرمول ۴.۵ نشان داده شده هیستوگرامهای فازی مورد بحث دردرس ۲ را معادل سازی می کند .

در **PFS** برای نمونه های طبقه بندی درهرطبقه **Ck** به صورت یک خروجی جدید نشان داده می شود پس فرمول ۴.۵ به صورت زیر هم نوشته می شود :

$$\hat{p}_{j,k} = \frac{\sum_{i=1}^{n} \overline{\mu}_A(x_i)\mu c_k(y_i)}{\sum_{i=1}^{n} \overline{\mu}_{A\,j}(x_i)} \qquad (۴.۶)$$

عملکرد مشخصه **XCk** بوسیله فرمول زیر نشان داده می شود :

$$x_{C_k}(y) = \begin{cases} 1 & if y=C_k \\ 0 & otherwise \end{cases} \qquad (۴.۷)$$

درتعریف این قسمت ،احتمالات آماری پارامترها تخمین زده می شوند . به PFSs درنمونه های طبقه بندی در تجزیه وتحلیل فرمولهای (۴.۵) و(۴.۶) در قسمت (۴.۱.۱) توجه می شود . همچنین در قسمت (۴.۱.۲) درنمونه های برگشتی PFSs بررسی می شود .

۴.۱.۱- نمونه های طبقه بندی درمسائل آماری :

دراین قسمت ثابت می شود که مسئله های احتمال که به وسیله فرمول (۴.۶) تخمین زده شده باشند غیر واقعی وناهماهنگ هستند وبا معیارهای ML سازگار نمی باشند .

همچنین کافی است یک عامل نمونه درفرمول(۴.۶) قرارداده شود تا غیر واقعی وناهماهنگ بودن تخمین های بدست آمده واینکه بیشینه سازی احتمال درست نمایی مجموعه اطلاعات انجام نمی شود اثبات گردد.

ملاحظه کنید که در **PFS** اگرمسئله طبقه بندی درخواست شده ۲ نوع باشد با **C1** و**C2** نمایش داده می شود . **PFS** یک ورودی **X=[0,1]** ویک قانون پایه شامل ۲ احتمال تئوری فازی دارد . در مقدمه **mfs** فازی **A1,A2** می نشیند پس خواهیم داشت :

$$\mu_{A_1}(x) = 1 - x \qquad and \qquad \mu_{A_2}(x) = x. \qquad (۴.۸)$$

دردنباله با توجه به فرمول (۳.۴) که **μ¯Aj=μAj** و **j=1,2** مفروض است که احتمال شرطی **C1** و**C2** برابر است با :

$$\Pr(C_1 \mid x) = 1 - x \qquad and \qquad \Pr(C_2 \mid x) = x. \qquad (۴.۹)$$

با استفاده ازفرمول (۳.۵) می توانیم احتمال های شرطی نا شناخته ای را که برای تخمین به **PFS** احتیاج ندارند ببینیم .

بااستفاده از فرمول (۴.۹) پارامترهای احتمال بدین صورت خواهند بود که : **P*1,1=P*2,2=1** و **P*1,2=P*2,1=0** (توجه کنید که در این مثال مقدمه **mfs** درفرمول

(۴.۸) به روشی انتخاب شده است که بدست آوردن تخمین درست احتمال شرطی PFS

را مشکل می نماید لذا بدست آوردن تخمین های درست احتمال شرطی پارامترهای احتمال

Pj,k نیزمشکل خواهد بود ودر نتیجه آنالیز تخمین های پارامترهای احتمالی ، غیرواقعی وناهماهنگ می باشد .

درادامه ۲قضیه که درارتباط باپارامترهای آماری فرمول (۴.۶) می باشد خواهد آمد . برای اثبات قضیه ها از مثال فوق استفاده میگردد .

قضیه(۴.۱):

برای نمونه های طبقه بندی شده در PFS بااستفاده از فرمول (۴.۶)اثبات کنید که تخمین های Pj,k ازپارامترهای احتمالی P*j,k غیرواقعی وناهماهنگ هستند .

اثبات : مثالی را که دربالا نشان داده شده ملاحظه نمایید . فرض کنید یک مجموعه اطلاعاتی شامل n نمونه طبقه بندی شده (i=1, … , n) (Xi,yi) برای تخمین پارامترهای احتمال درPFS دردسترس است . برای سادگی فرض کنید که X1, … ,Xn ارزشهای ثابتی دارند یعنی فقط Y1, … ,Yn نمونه هایی بارفتارهای متغیر هستند . برای مثال تخمین

P2,2 ازپارامتراحتمالی P*2,2 را ملاحظه کنید . ازفرمولهای (۴.۶) ،(۴.۷)

،(۴.۸)، (۴.۹)

چنین بدست می آید که :

$$Ep_{2,2} = E\left(\frac{\sum_{i=1}^{n} \bar{\mu}_{A_2}(x_i)\mu c_2(y_i)}{\sum_{i=1}^{n} \bar{\mu}_{A2}(x_i)}\right)$$

$$= \frac{\sum_{i=1}^{n} \bar{\mu}_{A2}(x_i)E(xc_2(y_i))}{\sum_{i=1}^{n} \bar{\mu}_{A2}(x_i)}$$

$$= \frac{\sum_{i=1}^{n} x_i(o(1-x_i)+1x_i)}{\sum_{i=1}^{n} x_i} \qquad (۱۰،۴)$$

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

حالا فرض کنید که(XiЄ(0,1 و(i=1,... ,n) سپس از فرمول (۴.۱۰) بدست

آورید که

(Ep2,2Є(0,1 تازمانیکه P*2,2=1 تخمین غیرواقعی از P2,2 باشد . این

بحث اعداد مستقلی از نمونه های طبقه بندی شده n راشامل میگردد.

همچنین از n→∞ تشکیل شده است .از دو مورد فوق نتیجه می شود که

تخمین P2,2 غیر واقعی و ناهماهنگ است .

معادله (۴.۶) تخمین های پایه رافقط وفقط برای اعدادمثبت Є .

lim Pr(|pj,k– Є)=1 n→∞ (۴.۱۱)

p*j,k|≤

دراینجا تخمین Pj,k ازیک مجموعه اطلاعات شامل n نمونه طبقه بندی شده بدست می آید .

این شرط می تواند همچنین به صورت Plim pj,k=p*j,k نوشته شود . شرط لازم برای

Plim pj,k=p*j,k این است که Epj,k=p*jk lim n→∞ باشد. (ببینید قضیه ۲.۹.۳۹ دردرس۱۲) تخمین pj,k از p*j,k باید واقعی و هماهنگ باشد .اگر چه فعلا اثبات شده که

Pj,k تخمین غیرواقعی وناهماهنگی از pj,k است . بنابراین نتیجه می شود که pj,k تخمین غیرواقعی از p*j,k می باشد واین کاملا تئوری ما را اثبات می کند .

قضیه ۴.۲ :

نمونه های طبقه بندی شده در PFS رادرنظر بگیرید یک مجموعه اطلاعاتی نیز داده شده است . پارامترهای احتمالی pj,k بااستفاده ازفرمول (۴.۶) تخمین زده شده اند واحتیاجی به بیشینه سازی احتمال درست نمایی مجموعه اطلاعات نمی باشد .

x	0.0	0.5	0.5	1.0
y	C1	C1	C2	C2

جدول ۴.۱: مجموعه اطلاعاتی که در اثبات قضیه ۴.۲ استفاده می شود .

اثبات : مثال داده شده در بالا را ملاحظه کنید . فرض کنید که مجموعه اطلاعاتی شامل ۴ نمونه طبقه بندی شده($i=1,2,3,4$)(X_i,y_i) برای تخمین پارامترهای احتمال در PFS دردسترس می با شد . مجموعه اطلاعات در جدول ۴.۱ نشان داده شده است .

نمونه های طبقه بندی شده را درفرمول (۴.۶) جایگزین کنید در نتیجه خواهیم داشت $P_{1,1}=P_{2,2}=0.75$ و $P_{1,2}=P_{2,1}=0.25$ سپس از فرمول (۳.۵)به دست می آید که

(۴.۱۲) $$\hat{p}(C_1|x)=0.75-0.5x \quad \text{و} \quad \hat{p}(C_2|x)=0.25+0.5x$$

احتمال درست نمایی مجموعه اطلاعات نشان داده می شود به وسیله

(۴.۱۳) $$L=\prod_{i=1}^{n} \hat{p}(Y_i|x_i)$$

حالا فرض کنید که نمونه های مجموعه اطلاعات مستقل ازهم می باشند برای پارامترهای احتمالی $P_{j,k}$ که به وسیله فرمول (۴.۶) تخمین زده شده اند با استفاده از فرمولهای (۴.۱۲) و (۴.۱۳) احتمال درست نمایی مجموعه اطلاعات جدول ۱-۴ برابرخواهد بود با

$9/64 \approx 0.14$. حالا ملاحظه کنید اگر پارامترهای احتمالی به

$P'_{1,1}=P'_{2,2}=1$ و $P'_{1,2}=P'_{2,1}=0$ تبدیل شوند با استفاده از فرمول

(۳.۵) نتیجه پارامترهای احتمالی برابر خواهد شد با:

$$P''(C2|x)=x \quad \text{و} \quad p''(C1|x)=1-x \qquad (۴.۱۴)$$

برای تبدیل پارامترهای احتمال $p'_{j,k}$ از فرمولهای (۴.۱۳) و (۴.۱۴) استفاده می شود که احتمال درست نمایی مجموعه اطلاعات در جدول ۱-۴ برابر با 0.25 خواهد شد . بنابراین تبدیل پارامترهای احتمال در ارزشهای بالاتر احتمال درست نمایی نتیجه بخش می باشد لذا پارامترهای احتمالی $P_{j,k}$ با استفاده از فرمول (۴.۶) تخمین زده می شوند . این مثال نشان می دهد که پارامترهای تخمین زده شده با استفاده از فرمول (۴.۶) احتیاج به بیشینه سازی احتمال درست نمایی مجموعه اطلاعات ندارند(واقعیت این است که مثال نشان می دهد که تبدیل پارامترهای احتمالی $P'_{j,k}$ احتمال درست نمایی مجموعه اطلاعات را بیشینه سازی می کند . ودرست است که تخمین ML پارامترهای احتمال دقیقا برابربا پارامترهای احتمالی $p^*_{j,k}$ به محض اتفاق افتادن مجموعه اطلاعات خاص در جدول ۱-۴ است .)

این اثبات قضیه را کامل می کند .

این موضوع توجه را جلب می کند که در یک سیستم که ورودی **x** به روشهای جدید تقسیم می شود(x) **i.e.μ¯Aj** برابرخواهد بودبرای **a , ... ,j=1** وبرای همه (**x Є X**) با**0** یا **1**.

برای پارامترهای احتمال که با استفاده از فرمول (۴.۶) تخمین زده می شوند میتوانیم واقعی وهماهنگ بودن با معیارهای **ML** را نشان دهیم . (در این قسمت اثبات نمی شود)

بنابراین در سیستم های جدید تخمین پارامترهابا مطلوبیت آماری ممکن است با تخمین هر پارامتر به تفکیک وبا استفاده از فرمول (۴.۶) بدست آید . در یک سیستم فازی اگر چه با استفاده از قضایای ۴.۲ و۴.۱ تخمین پارامترها با مطلوبیت آماری با تخمین هرپارامتر به تفکیک وبا استفاده از فرمول (۴.۶) بدست نمی آید در عوض پارامترها در یک سیستم فازی می توانند به طور همزمان تخمین زده شوند واین به پیشنهاد مطرح شده در بخش ۴.۲ نزدیک است .

۴.۱.۲-احتمال آماری در مسئله های برگشتی :

دراین قسمت اثبات خواهیم کردکه تخمین پارامترهای شرطی بااستفاده ازفرمول (۴.۵) غیرواقعی است وبامعیارهای **ML** هماهنگی و سازگاری ندارد .

برای اثبات کافی است که یک عامل به عنوان مثال درفرمول (۴.۵) قرارداده شود

تا نشان دهد تخمین هایی که غیرواقعی وناهماهنگ می باشند وهمچنین اثبات شود که بیشینه سازی مجموعه اطلاعات دردسترس انجام نمی شود .

باید توجه شود که این قسمت خیلی مشابه قسمت قبل می باشد . تنها فرق موجود این است که این قسمت درارتباط با **PFSs** برای نمونه های برگشتی در عوض **PFSs** برای نمونه های طبقه بندی است . درنظر داشته باشید که **PFS** یک راه کاربردی در مسائل برگشتی است دراینگونه مسائل **PFS** یک ورودی **x=[0,1]** ویک خروجی **y=[0,1]** دارد . اساس این سیستم ۲ احتمال قانون فازی را شامل می شود . در مقدمه **mfs** فازی **A1** و **A2** بوسیله فرمول (۴.۱۵) نشان داده می شود .

$$\mu A1(x)=1-x \quad و \quad \mu A2(x)=x \qquad (۴.۱۵)$$

از فرمول (۳.۴) نتیجه می شود که : $\mu^{-}Aj=\mu Aj$ و $j=1,2$. خروجی **y** بااستفاده از مجموعه فازی به **C1** و **C2** تقسیم می شود . **mfs** ازاین مجموعه فلزی بااستفاده از فرمول (۴.۱۶) به دست می آید .

$$\mu C1(y)=1-y \quad و \quad \mu C2(y)=y \qquad (۴.۱۶)$$

توجه کنید که فرمول (۳.۱۱) شرط کافی است یعنی اینکه **y** باید خوب تعریف شده باشد . اگر فرض کنیم که **pdf** شرطی **y** به شرط **x** برابر باشد با

$P(y|x)=4xy-2x-2y+2$

این **pdf** شرطی نشان می دهد که به تخمین **PFS** احتیاج داریم . باستفاده از فرمولهای

(۳.۵) ، (۳.۱۲) ،(۳.۱۳) می توانیم ببینیم که در یک **PFS** که به طور صحیح تخمین زده شده باشد **pdf** شرطی از فرمول (۴.۱۷) بدست می آید . پارامترهای احتمالی بدست آمده عبارتند از $P^*1,1=P^*2,2=1$ و $P^*1,2=P^*2,1=0$ (توجه کنید که دراین مثال مقدمه **mfs** درفرمول (۴.۱۵) ونتیجه **mfs** در فرمول (۴.۱۶) به روشی می باشد که شامل **PFS** که **pdf** شرطی را به طور صحیح در فرمول (۴.۱۷) تخمین زده باشد نیز می شود.

ممکن است آن **PFS** را که **pdf** شرطی را به طور صحیح تخمین زده باشدشامل نشود لذا پارامترهای شرطی **p*jk** ممکن است صحیح نباشند ونتیجه آن نیز ممکن است تجزیه وتحلیل واقعی وهماهنگی از تخمین پارامترهای احتمال نداشته باشد .

دردنباله برای احتمال آماری ۴.۵ دو قضیه مورد توجه می باشد . قضیه ها را با استفاده از مثال فوق اثبات خواهیم کرد .

قضیه ۴.۳:

با استفاده از فرمول (۴.۵) در یک ***PFS*** برای نمونه های برگشتی اثبات کنید که تخمین های ***Pj,k*** از پارامترهای شرطی ***p*jk*** غیر واقعی و ناهماهنگ می باشند .

x	0.0	0.5	1.0
y	0.0	0.5	1.0

جدول ۴.۲: مجموعه اطلاعاتی که در اثبات قضیه ۴-۴ استفاده می شوند

اثبات : مثال نشان داده شده در بالا را ملاحظه کنید . فرض کنید مجموعه اطلاعاتی شامل

n نمونه((i=1,...,n (xi,yi) برای تخمین پارامترهای احتمالی درPFS دردسترس می باشد برای سادگی فرض کنید که X1, ... ,Xn ارزشهای ثابتی دارند یعنی فقط Y1,...,Yn نمونه هایی بارفتارهای متغیر هستند . برای مثال تخمین P2,2 ازپارامتراحتمالی P*2,2 را ملاحظه کنید. ازفرمولهای (۴.۵)،(۴.۱۵)،(۴.۱۶)،(۴.۱۷) چنین بدست می آید که :

$$Ep_{2,2} = E\left(\frac{\sum_{i=1}^{n} \bar{\mu}_{A_2}(x_i)\mu c_2(y_i)}{\sum_{i=1}^{n} \bar{\mu}_{A2}(x_i)}\right) \quad :(۴.۱۸)$$

$$= \frac{\sum_{i=1}^{n} \bar{\mu}_{A2}(x_i) E(xc_2(y_i))}{\sum_{i=1}^{n} \bar{\mu}_{A2}(x_i)}$$

$$= \frac{\sum_{i=1}^{n} x_i \int_0^1 y p(y \mid x_i) dy}{\sum_{i=1}^{n} x_i}$$

$$= \frac{\sum_{i=1}^{n} x_i (\frac{1}{3} x_i + \frac{1}{3})}{\sum_{i=1}^{n} x_i}$$

زیرا(XiЄ(0,1 و (i=1,... ,n) ودردنباله از فرمول (۴.۱۸) نتیجه می شود

که:

Ep2, 2 Є همچنین P*2,2= 1 وتخمین P2,2 غیرواقعی است . این بحث

دراعداد ونمونه

های n دارای ظرفیت مستقلی می باشد بنابراین برای n→∞ هم ظرفیت دارد

.

ازاینها نتیجه می شود که تخمین P2,2 همچنین غیر واقعی است . معادله

(۴.۵) اثبات تخمینها را پیگیری می نماید اگروفقط اگر Plimpj,k=p*j,k .

شرط لازم برای Plimpj,k=p*j,k این است که lim n→∞ و

Epj,k=p*j,k (به قضیه ۳۹-۹-۲درس ۱۲ توجه کنید . تخمین pj,k از p*j,k

باید مجانب وواقعی باشد . اگرچه بزودی اثبات میگردد که pj,k تخمین

غیرواقعی مجانبی از p*j,k می باشد . بنابراین غیرواقعی بودن تخمین pj,k

اثبات می شود که همان اثبات قضیه ۴.۳ می باشد .

قضیه ۴.۴:

یک PFS ازنمونه های برگشتی را در نظر بگیرید نشان دهید در یک مجموعه اطلاعاتی پارامترهای احتمالی **pj,k به وسیله فرمول (۴.۵) تخمین زده می شوند وبه بیشینه سازی احتمال درست نمایی مجموعه اطلاعات احتیاج نمی باشد .**

اثبات :

مثالی را که در بالا اورده شده ملاحظه کنید . فرض کنید مجموعه اطلاعاتی شامل ۳ نمونه (xi,yi) (i=1,2,3) ازتخمین پارامترهای احتمال در PFS دردسترس می باشد. مجموعه اطلاعات جدول ۴.۲ را در فرمول (۴.۵) جایکزین کنید نتیجه این است که p1,1=p2,2=5/6

و p1,2=p2,1=1/6 سپس از فرمولهای (۳.۵)،(۳.۱۲)،(۳.۱۳) بدست می آید که :

$$p\hat{}(x|y)=1/3(8xy-4x-4y+5) \qquad (۴.۱۹)$$

احتمال درست نمایی مجموعه اطلاعات به وسیله فرمول (۴.۱۳) بدست می آید . برای تخمین پارامترهای احتمالی pj,k از فرمول (۴.۵) وبه دنبال آن از فرمولهای (۴.۱۳) و

(۴.۱۸) استفاده می شود که احتمال درست نمایی مجموعه اطلاعات در جدول ۴-۲ برابراست با 25/9≈2/78 حالا تبدیل پارامترهای احتمال

$p'_{1,1}=p'_{2,2}=1$ و $p'_{1,2}=p'_{2,1}=0$ را با استفاده از فرمول (۳.۵) ملاحظه نمایید :

$$\hat{p}(x|y)=4xy-2x-2y+2 \qquad (4.20)$$

برای تبدیل پارامترهای احتمالی $p'_{j,k}$ با توجه به فرمول (۴.۱۳) و (۴.۲۰) احتمال درست نمایی مجموعه اطلاعات در جدول ۲-۴ برابر با ۴ خواهد بود . بنابراین تبدیل پارامترهای احتمالی نتیجه ای است از احتمال درست نمایی ارزشهای بالا سپس پارامترهای احتمالی $p_{j,k}$ با استفاده از فرمول (۴.۵) تخمین زده می شود .

این مثال نشان می دهد که پارامترهای احتمالی تخمین زده شده با استفاده از فرمول (۴.۵) احتیاج به بیشینه سازی احتمال درست نمایی مجموعه اطلاعات ندارد . (در واقع در مثال نشان داده شده است که تبدیل پارامترهای احتمالی $p'_{j,k}$احتمال درست نمایی مجموعه اطلاعات را بیشینه سازی می کند . البته ML پارامترهای احتمال را دقیقا برابربا پارامترهای احتمال $p^*_{j,k}$ تخمین می زند به محض اینکه نتیجه ای معین از مجموعه اطلاعات جدول ۲-۴ اتفاق افتد .) واین اثبات قضیه را کامل می نماید .

توجه کنید که در یک سیستم که ورودی x وخروجی y به روش جدید تقسیم بندی می شوند $\mu^{-}A_j(x)$ برای همه($x\in X$)($j=1,\ldots,a$) برابر خواهد بود با صفر و یک و$\mu cL(y)$ برای همه $(k=1,\ldots,c)(y\in Y)$ نیز برابر خواهد بود با

صفر و یک . پارامترهای احتمال تخمین زده شده با استفاده از فرمول (۴.۵) واقعی هستند وبا معیارهای ML هماهنگی و سازگاری دارند . (اینجا اثبات نمی شود) بنابراین در سیستم جدید ممکن است پارامترهای تخمین زده شده با آمارهای احتمالی مطلوب بااستفاده از فرمول (۴.۵) وبه وسیله تخمین زدن هر پارامتر به تفکیک بدست آید . در یک سیستم فازی با توجه به قضیه های ۳-۴ و ۴-۴ پارامترهای تخمین زده شده با آمارهای احتمالی مطلوب بااستفاده از فرمول (۴.۵) وبه وسیله تخمین زدن هر پارامتر به تفکیک بدست نمی آید در عوض پارامترها در سیستم فازی به طور همزمان تخمین زده می شوند وبدین گونه به پیشنهاد مطرح شده در قسمت بعدی نزدیک می شویم .

۴-۲ : روش افزایش احتمال درست نمایی

در این قسمت پیشنهاد می کنم از معیار های ML برای تخمین پارامترهای احتمال در PFS استفاده شود برعکس روش احتمال شرطی که در بخش قبلی مورد بحث قرار گرفت همه پارامترها ی احتمال درPFS به طور همزمان دریک فاصله نزدیک تخمین زده می شوند .

فرض کنید دراین قسمت که هم مقدمه وهم نتیجه mfs درPFS مشخص است و نیازی به بیشینه سازی بیشتر نیست . (برای مسئله های طبقه بندی هم تخمین ML وهم mfs در پارامترهای احتمالی در قسمت ۲-۵ بررسی می شود) باید توجه شود که تمرکز این قسمت روی PFSs برای تخمین شرطی

pdfs می باشد . بنابراین نتیجه این بخش همچنین به **PFSs** برای نمونه های طبقه بندی ارتباط خواهد داشت تا زمانیکه یک**PFS** برای نمونه های طبقه بندی بتواند مورد مخصوصی ازیک **PFS** برای تخمین شرطی **pdfs** باشد .

توجه کنید که تخمین **ML** پارامترهای احتمال دردرس ۲۶ مورد بحث قرار می گیرد .

همچنین توجه کنید که که با جایگزین کردن فرمول (۳.۵) در فرمول (۳.۱۲) نتیجه می شود که :

$$\hat{p}(y \mid x) = \sum_{k=1}^{c} p(y \mid C_K) \sum_{j=1}^{a} \overline{\mu} A_j(x) p_{j,k} \quad (۴.۲۱)$$

پارامترهای احتمال **pj,k** درفرمول (۴.۲۱) باید شرطهای فرمول (۳.۲) و (۳.۳) را برآورده کند . از فرمول (۳.۳) بدست می آید که**pj,c=Σ c-1,k=1pj,k**

بنابراین فرمول

(۴.۲۱) می تواند همچنین به صورت زیر نوشته شود :

$$\hat{p}(y \mid x) = \Big(\sum_{k=1}^{c-1} p(y \mid C_K) \sum_{j=1}^{a} \overline{\mu} A_j(x) p_{j,k}\Big) + \quad (۴.۲۲)$$
$$p(y \mid C_c) \sum_{j=1}^{a} \overline{\mu} A_j(x) \Big(1 - \sum_{k=1}^{c} p_{j,k}\Big)$$

پارامترهای احتمالی **pj,k** در فرمول (۴.۲۲) باید برآورده کند :

$p_{j,k} \geq 0$ برای $j=1,\ldots,a$ و $k=1,\ldots,c-1$ (۴.۲۳)

$$\sum_{k=1}^{c} p_{j,k} \leq 1 \qquad for\, j = 1,\ldots,a \tag{۴.۲۴}$$

حالا فرض کنید که مجموعه اطلاعات شامل n نمونه $(x_i,y_i)(i=1,\ldots,n)$ برای تخمین پارامترهای احتمال PFS دردسترس است . پارامترهای اصلی $P=[p_{j,k}](j=1,\ldots,a)$ را جایگزین نمایید احتمال درست نمایی مجموعه اطلاعات برابر خواهد بود با :

$$L(P)=\prod_{i=1}^{n} \hat{p}(Y_i|x_i) \tag{۴.۲۵}$$

حالا فرض کنید که نمونه های موجود در مجموعه اطلاعات مستقل از هم باشند . احتمال شرطی در p باید شرطهای فرمول های (۴.۲۳) و (۴.۲۴) را برآورده کند . بااستفاده از فرمول (۴.۲۲) و (۴.۲۵) لگاریتم احتمال درست نمایی می تواند به صورت زیر نوشته شود :

$$\begin{aligned} l(p) &= in\left(\prod_{i=1}^{n} \hat{p}(y_i \mid x_i)\right) \\ &= \sum_{i=1}^{n} 1n\hat{p}(y_i \mid x_i) \\ &= \sum_{i=1}^{n} 1n\left(\left(\sum_{k=1}^{c-1} p(y \mid C_K)\sum_{j=1}^{a} \overline{\mu A_j}(x) p_{j,k}\right.\right. \\ &\quad \left.\left. + p(y_i C_c)\sum_{j=1}^{a} \overline{\mu A_j}(x_i)\left(1 - \sum_{k=1}^{c-1} p_{j,k}\right)\right)\right) \end{aligned} \tag{۴.۲۶}$$

تخمین ML ازپارامترهای احتمالی PFS به وسیله بیشینه سازی لگاریتم احتمال درست نمایی عملکرد L(P) در فرمول (۴.۲۶) با در نظر گرفتن پارامتر اصلی P برآورد می شود

ازبیشینه سازی L(P) به وسیله شرطهای فرمول (۴.۲۳) و (۴.۲۴) جلوگیری می شود .

تا زمانیکه عملکرد L(P) غیر خطی است پیدا کردن تخمین های ML پارامترهای شرطی یک مسئله برنامه ریزی شده غیر خطی است . درادامه قضیه را ملاحظه کنید .

قضیه ۵-۴ :

لگاریتم احتمال درست نمایی عملکرد L(P) در فرمول (۴.۲۶) واگراست .

اثبات :

زمانیکه جمع عملگرهای واگرا ، واگراست کافی است اثبات کنیم که :

$$\begin{aligned}\phi(p) &= in\left(\prod_{i=1}^{n} \hat{p}(y_i \mid x_i)\right) \\ &= \sum_{i=1}^{n} 1n\hat{p}(y_i \mid x_i) \\ &= \sum_{i=1}^{n} 1n\left(\left(\sum_{k=1}^{c-1} p(y \mid C_K)\sum_{j=1}^{a} \overline{\mu A_j}(x) p_{j,k}\right.\right. \\ &\left.\left. + p(y_i C_c)\sum_{j=1}^{a} \overline{\mu A_j}(x_i)\left(1-\sum_{k=1}^{c-1} p_{j,k}\right)\right)\right)\end{aligned} \quad (۴.۲۷)$$

عملکرد واگرا برای همه(xЄX) وهمه (yЄY) است توجه کنید که Φ(p) از هر پارامتر اصلی p برای p^(y|x)>0 بدست می آید . علاوه براین Φ(p) در دو مورد پیوسته متفاوت است .

حالت اول وحالت دوم مشتق نسبی Φ(p) هستند که نشان داده می شوند به وسیله :

$$\frac{\partial\phi(P)}{\partial p_{\alpha,\tau}}=\frac{\mu A_{\alpha}(x)(p(y\,|\,C_{\tau})-p(y\,|\,C_{c}))}{\hat{p}(y\,|\,x)} \quad (۴.۲۸)$$

(۴.۲۹)

$$\frac{\partial^{2}\phi(P)}{\partial p_{\alpha,\tau}\partial p_{\beta,\delta}}=-\frac{\bar{\mu}A_{\alpha}(x)(p(y\,|\,C_{\tau})-p(y\,|\,C_{c}))\bar{\mu}A_{\beta}(x)(p(y\,|\,C_{\delta})-p(y\,|\,C_{c}))}{\hat{p}(y\,|\,x)^{2}}$$

در جاهایی از فرمول (۴.۲۹) که α,β=1,...,a و γ,σ=1,...,c-1 هر پارامتر اصلی p برای هر Φ(p) بدست می آید از :

$$\frac{\partial^2 \phi(P)}{\partial p^2_{\alpha,\tau}} = -\left(\frac{\bar{\mu} A_\alpha(x)(p(y|C_\tau)-p(y|C_c))}{\hat{p}(y|x)}\right)^2 \leq 0 \qquad (۴.۳۰)$$

اجازه دهید Dm بر m*m دلالت کند وتعیین شود به وسیله :

$$D_m = \begin{vmatrix} \frac{\partial^2 \phi(P)}{\partial p_{\alpha_1,\tau_1}\partial p_{\beta_1,\delta_1}} & \cdots & \frac{\partial^2 \phi(P)}{\partial p_{\alpha_1,\tau_1}\partial p_{\beta_m,\delta_m}} \\ \frac{\partial^2 \phi(P)}{\partial p_{\alpha_m,\tau_m}\partial p_{\beta_1,\delta_1}} & \cdots & \frac{\partial^2 \phi(P)}{\partial p_{\alpha_m,\tau_m}\partial p_{\beta_m,\delta_m}} \end{vmatrix} \qquad (۴.۳۱)$$

در اینجا برای q≠r و(αq,γq)≠(αr ,γr) و (βq,σq)≠(βr,σr)

ازفرمول(۴.۲۹) استفاده می شود که نشان می دهد برای تعیین هر Dm با

m=2 وبرای هر پارامتر اصلی p ,Φ(p) بدست می آید از :

$$D_2 = \frac{\partial^2 \phi(P)}{\partial p_{\alpha_1,\tau_1}\partial p_{\beta_2,\delta_2}} \frac{\partial^2 \phi(P)}{\partial p_{\alpha_2,\tau_2}\partial p_{\beta_2,\delta_2}} - \frac{\partial^2 \phi(P)}{\partial p_{\alpha_2,\tau_2}\partial p_{\beta_1,\delta_1}} \frac{\partial^2 \phi(P)}{\partial p_{\alpha_1,\tau_1}\partial p_{\beta_2,\delta_2}} = 0 \qquad (۴.۳۲)$$

بوسیله ارتباط مراحل تعیین وگسترش انها فرمول (۴.۳۲) بدست می آید که در آن نیز **Dm** برای همه مشخصه ها یی با **m>2** برابر صفر خواهد بود .

همچنین با ترکیب آن فرمول

(۴.۳۰) نتیجه می شود که پارامتر اصلی **Φ(p),** تقریبا در همه جا منفی است (مراجعه به قضیه (۱۱-E – ۱) در درس ۳۲) در نتیجه,**Φ(p)** یک عملگر واگراست . واین اثبات قضیه را کامل می کند.

توجه داشته باشید که این قضیه عمومی است یعنی نمی توانیم درباره **μAj** **mfs** در**PFS**

تصوری داشته باشیم . در مسائل برنامه ای غیر خطی تخمین های **ML** پارامترهای احتمالی **PFS** را پیدا می کنند عملگرها در محدودیتها به وسیله فرمول های (۴.۲۳) و

(۴.۲۴) که همگی خطی هستند نشان داده می شوند . از همه اینها نتیجه می شود که این عملگرها همگرا هستند .زمانی برحسب تئوری ۵-۴ موضوع عملگرها مسائل برنامه ای غیر خطی است که در حقیقت همان مسائل برنامه ای همگرا می باشند . مسائل برنامه ای همگرا ویژگیهای دردسترسی دارند که عبارت است از روش مساعد محلی و روش مساعد کلی بنابراین پیدا کردن یک روش مساعد کلی به عنوان یک راه حل باید نسبتا آسان باشد .