محاسبه اطلاع فیشر در هر مجموعه دلخواه از آماره‌های مرتب

مهدی روزبه[۱] سیّد محمّد مهدی طباطبایی[۲]

چکیده

وقتی n مشاهده‌ی مستقل و هم توزیع داریم، یک سؤال جالب این است که چگونه اطلاع فیشر در بین آماره‌های مرتب توزیع شده است. دستور کلی برای محاسبه اطلاع فیشر در آماره‌های مرتب ساده است ولی جزئیات محاسبه‌ی آن پیچیده می‌باشد. یک روش غیرمستقیم استفاده از تجزیه‌ی اطلاع فیشر در آماره‌های مرتب است که محاسبات را ساده می‌کند. برخی معادلات بازگشتی برای اطلاع فیشر در آماره‌های مرتب استخراج شده که به کمک آن‌ها محاسبات ساده می‌شوند. اطلاع فیشر در نخستین r آماره‌ی مرتب یک انتگرال r گانه می‌باشد که به کمک این معادلات به انتگرال‌های دوگانه و معمولی ساده می‌شود. هم‌چنین این نتایج ما را قادر می‌سازد که اطلاع فیشر در هر مجموعه از آماره‌های مرتب را به سادگی محاسبه نماییم. در انتها کاربرد آنها را در مسائل فضای بهینه مورد بررسی قرارمی‌دهیم.

واژه‌های کلیدی: آماره‌ی مرتب، اصل والد، اطلاع فیشر، اطلاع فیشر تعمیم یافته، اطلاع فیشر حاشیه‌ای، چندک نمونه‌ای، رابطه‌ی بازگشتی، فضای بهینه.

۱ مقدمه

فرض کنید یک نمونه‌ی تصادفی به حجم n از توزیعی پیوسته با تابع توزیع تجمعی $F(x;\theta)$ و تابع چگالی احتمال $f(x;\theta)$ باشد، که θ یک پارامتر با مقدار حقیقی است. هم‌چنین برای $f(x;\theta)$ شرایط نظم برقرار بوده و تابعی روی $\chi \times \Theta \longmapsto R^{1}$ باشد که در آن χ فضای نمونه برای متغیر تصادفی X و Θ فضای پارامتر است. این سؤال که کدام قسمت از نمونه‌ی مرتب شده شامل اطلاع بیشتری است توسط توکی [۱۶] مورد بحث قرار گرفته است. در این بخش محاسبه‌ی اطلاع فیشر را که نقش مهمی در آمار استنباطی و برآوردیابی بازی می‌کند مورد بررسی قرار می‌دهیم.

فرض کنید $X_{(r:n)}$، r امین آماره‌ی مرتب با تابع چگالی احتمال $f_{r:n}$ و k امین گشتاور $\mu_{r:n}^{(k)}$ باشد. اطلاع فیشر درباره‌ی پارامتر θ در مجموعه‌ی اولین r آماره‌ی مرتب $(X_{(1:n)},...,X_{(r:n)})$ به صورت زیر تعریف می‌شود:

$$I_{1...r:n}(\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{X_{(r-1:n)}} \cdots \int_{-\infty}^{X_{(2:n)}} \left(\frac{\partial}{\partial\theta}\log f_{1...r:n}\right)^{2} dF_{1...r:n} \qquad (1)$$

[۱] دانشگاه فردوسی مشهد
[۲] دانشگاه فردوسی مشهد

به طوری که $f_{1 \ldots r:n}$ و $F_{1 \ldots r:n}$، به ترتیب توابع چگالی احتمال و توزیع تجمعی $(X_{(1:n)}, \ldots X_{(r:n)})$ می‌باشد. در معادله‌ی فوق، محاسبه‌ی $I_{1 \ldots r:n}(\theta)$ پیچیده است زیرا شامل انتگرال روی r متغیر تصادفی است.

اما به جای معادله‌ی (۱) می توانیم تجزیه‌ی اطلاع فیشر در آماره‌های مرتب را به شکل زیر مورد بررسی قرار دهیم: (خاصیت مارکفی آماره‌های مرتب)

$$I_{1 \ldots n:n}(\theta) = I_{1 \ldots r:n}(\theta) + I_{r+1 \ldots n|r:n}(\theta) \qquad (۲)$$

به طوری که $I_{r+1 \ldots n|r:n}(\theta)$ میانگین اطلاع شرطی در $(X_{(r+1:n)}, \ldots, X_{(n:n)})$ با شرط $X_{(r:n)} = x_{(r:n)}$ می‌باشد. در معادله‌ی اخیر به محاسبه‌ی $I_{1 \ldots r:n}(\theta)$ علاقه‌مندیم، اما برای بدست آوردن این رابطه ابتدا باید $I_{r+1 \ldots n|r:n}(\theta)$ را مورد بررسی قرار دهیم و سپس مقدار $I_{1 \ldots r:n}(\theta)$ را از معادله‌ی (۲) بدست آوریم. این رویکرد را با مطالعه‌ی اطلاع فیشر در آماره‌های مرتب تعمیم می‌دهیم. رابطه‌ی بازگشتی بین گشتاورهای آماره‌های مرتب برای هر دو توزیع پیوسته و گسسته توسط چندین مؤلف مورد مطالعه قرار گرفته است. معادله‌ی بازگشتی استاندارد ابتدا توسط کل [۵] به شکل زیر به دست آمد:

$$\mu_{r:n-1}^{(k)} = (n-r)\mu_{r:n}^{(k)} + r\mu_{r+1:n}^{(k)}$$

که این نتیجه به طور مستقیم از روابط بازگشتی مربوط بین توابع چگالی احتمال آماره‌های مرتب استخراج شده است. روابط بازگشتی فوق بین توابع چگالی احتمال آماره‌های مرتب فقط به منظور بدست آوردن روابط بازگشتی بین گشتاورهای آماره‌های مرتب مورد بررسی قرار گرفتند. اما در اینجا به کمک آنها روابط بازگشتی برای اطلاع فیشر در آماره‌های مرتب استخراج می‌کنیم که محاسبه‌ی اطلاع فیشر در آماره‌های مرتب به کمک آنها ساده می‌شود.

هم‌چنین نتایج ساده‌ای را بدست می آوریم که اطلاع فیشر در هر مجموعه از آماره‌های مرتب می‌تواند به سادگی به صورت ترکیب خطی از اطلاع فیشر در زوج‌های آماره‌های مرتب نمایش داده شود. بنابراین با محاسبه‌ی اطلاع فیشر در تمام مجموعه‌های دوتایی از آماره‌های مرتب به سادگی می‌توان اطلاع فیشر در هر مجموعه‌ی دلخواه از آماره‌های مرتب را محاسبه نمود.

در پایان کاربرد این معادلات را در مسائل فضای بهینه مورد بررسی قرار می‌دهیم.

۲ تجزیه‌ی اطلاع فیشر در آماره‌های مرتب

چون اطلاع فیشر دارای خاصیت جمع‌پذیری به عنوان یک اندازه از اطلاعات است رابطه‌ی (۲) به وسیله‌ی خاصیت مارکفی آماره‌های مرتب به دست می آید و چون $f_{r+1 \ldots n|r:n}$ تابع چگالی توام آماره‌های مرتب در یک نمونه به حجم $(n-r)$ از توزیع اصلی که از سمت چپ در نقطه‌ی $x = x_{(r:n)}$ بریده شده است، می‌باشد. به وسیله‌ی اصل والد (دیوید [۶]، $I_{r+1 \ldots n|r:n}(\theta)$ می‌تواند به صورت زیر نوشته شود:

$$I_{r+1 \ldots n|r:n}(\theta) = (n-r)\int_{-\infty}^{\infty} g(\omega;\theta) f_{r:n}(\omega;\theta) d\omega \qquad (۳)$$

به طوری که

$$\begin{aligned} g(\omega;\theta) &= \int_{\omega}^{\infty}\left\{\frac{\partial}{\partial\theta}\log\frac{f(x;\theta)}{1-F(x;\theta)}\right\}^{2} \\ &\times \frac{f(x;\theta)}{1-F(x;\theta)}dx \end{aligned}$$

بنابراین می‌توان انتگرال r گانه $I_{1\ldots r:n}(\theta)$ را بر حسب انتگرال دو گانه‌ی $nI_{1:1}(\theta) - I_{r+1\ldots n|r:n}(\theta)$ نوشت. و چون هدف ما در معادله‌ی (۲) محاسبه‌ی $I_{1\ldots r:n}(\theta)$ است، فقط کافی است $I_{r+1\ldots n|r:n}(\theta)$ را محاسبه نماییم زیرا همیشه راجع به $I_{1\ldots :n}(\theta)$ اطلاع داریم. به عنوان مثال مقدار $I_{r+1\ldots n|r:n}(\theta)$ برای توزیع نمایی با پارامتر مقیاس با استفاده از خاصیت فقدان حافظه برابر با $\frac{n-r}{\theta^{2}}$ است، پس می‌توان نوشت:

$$I_{1\ldots r:n}(\theta) = \frac{n}{\theta^{2}} - \frac{n-r}{\theta^{2}} = \frac{r}{\theta^{2}}$$

۳ روابط بازگشتی

رابطه‌ی ۳–۱

$$\begin{aligned} I_{r\ldots s:n-1}(\theta) &= \frac{n-s-1}{n}I_{r\ldots s:n}(\theta) + \frac{s-r+1}{n} \\ &\times I_{r\ldots s+1:n}(\theta) + \frac{r-1}{n}I_{r+1\ldots s+1:n}(\theta) \end{aligned} \qquad (4)$$

این رابطه، رابطه‌ی بازگشتی دیگری را هم برای اطلاع فیشر در نخستین r آماره‌ی مرتب شامل می‌شود:

$$\begin{aligned} I_{1\ldots r:n-1}(\theta) &= \\ \frac{n-r-1}{n} & I_{1\ldots r:n}(\theta) + \frac{r}{n}I_{1\ldots r+1:n}(\theta) \end{aligned} \qquad (5)$$

برهان: با استفاده از رابطه (۳) می‌توان نوشت:

$$I_{r+1\ldots n-1|r:n-1}(\theta) = (n-r-1)\int_{-\infty}^{\infty} g(\omega;\theta)f_{r:n-1}(\omega;\theta)d\omega$$

اکنون با استفاده از رابطه‌ی

$$f_{r:n-1}(\omega;\theta) = \frac{n-r}{n}f_{r:n}(\omega;\theta) + \frac{r}{n}f_{r+1:n}(\omega;\theta)$$

که اثبات آن در دیوید [٦] می‌باشد، می‌توان نوشت:

$$\begin{aligned} I_{r+1\ldots n-1|r:n-1}(\theta) &= \tfrac{n-r-1}{n} \\ &\times\left\{(n-r)\int_{-\infty}^{\infty} g(\omega;\theta)f_{r:n}(\omega;\theta)\,d\omega\right\} \\ &+\tfrac{r}{n}\left\{(n-r-1)\int_{-\infty}^{\infty} g(\omega;\theta)f_{r+1:n}(\omega;\theta)d\omega\right\} \\ &= \tfrac{n-r-1}{n}I_{r+1\ldots n|r:n}(\theta) + \tfrac{r}{n}I_{r+2\ldots n|r+1:n}(\theta) \end{aligned}$$

بنابراین با استفاده از رابطه‌ی (۲) به نتیجه‌ی زیر می‌رسیم:

$$I_{1\ldots r:n-1}(\theta) = \frac{n-r-1}{n}I_{1\ldots r:n}(\theta) + \frac{r}{n}I_{1\ldots r+1:n}(\theta)$$

بطور مشابه برای $I_{s\ldots n-1:n-1}(\theta)$ می‌توان نوشت:

$$\begin{aligned} I_{s\ldots n-1:n-1}(\theta) &= \frac{n-s}{n}I_{s\ldots n:n}(\theta) \\ &+ \frac{s-1}{n}I_{s+1\ldots n:n}(\theta) \end{aligned} \qquad (6)$$

اکنون با استفاده از رابطه (۲) می‌توان نوشت:

$$I_{s+1\ldots n|s:n}(\theta) = I_{r\ldots s:n}(\theta) - I_{r\ldots n:n}(\theta) \qquad (7)$$

حال با استفاده از روابط (۷) و

$$I_{Y|X}(\theta) = I_{Y,X}(\theta) - I_X(\theta)$$

می‌توان نوشت:

$$I_{r\ldots s:n}(\theta) = I_{1\ldots s:n}(\theta) + I_{r\ldots n:n}(\theta) - I_{1\ldots n:n}(\theta) \quad (8)$$

اکنون با استفاده از روابط (۵) و (۶) و (۸) می‌توان نوشت:

$$\begin{aligned}
& I_{r\ldots s:n-1}(\theta) = I_{1\ldots s:n-1}(\theta) + I_{r\ldots n-1:n-1}(\theta) \\
- \;& I_{1\ldots n-1:n-1}(\theta) = \frac{n-s-1}{n} I_{1\ldots s:n}(\theta) \\
+ \;& \frac{s}{n} I_{1\ldots s+1} + \frac{n-r}{n} I_{r\ldots n:n}(\theta) \\
+ \;& \frac{r-1}{n} I_{r+1\ldots n:n}(\theta) - \frac{n-1}{n} I_{1\ldots n:n}(\theta) \\
= \;& \frac{n-s-1}{n}[I_{r\ldots s:n}(\theta) - I_{r\ldots n:n}(\theta) + I_{1\ldots n:n}(\theta)] \\
+ \;& \frac{s}{n}[I_{r\ldots s+1:n} - I_{r\ldots n:n} + I_{1\ldots n:n}(\theta)] \\
+ \;& \frac{n-r}{n}[I_{r\ldots s+1:n}(\theta) + I_{s+2\ldots n|s+1:n}(\theta)] \\
+ \;& \frac{r-1}{n}[I_{r+1\ldots s+1:n}(\theta) + I_{s+2\ldots n|s+1:n}(\theta)] \\
- \;& \frac{n-1}{n} I_{1\ldots n:n}(\theta) = \frac{n-s-1}{n} I_{r\ldots s:n}(\theta) \\
+ \;& \frac{s-r+1}{n} I_{r\ldots s+1:n}(\theta) + \frac{r-1}{n} I_{r+1\ldots s+1:n}.
\end{aligned}$$

رابطه‌ی ۳-۲ رابطه‌ی بازگشتی بین اطلاع فیشر مجموعه‌ی دوتایی از آماره‌های مرتب به شکل زیر است:

$$\begin{aligned}
I_{r\ldots s:n-1}(\theta) &= \frac{n-s}{n} I_{rs:n}(\theta) \\
&+ \frac{s-r-1}{n} I_{rs+1:n}(\theta) + \frac{r}{n} I_{r+1s+1:n}(\theta) \\
&+ \frac{1}{n}[I_{r\ldots s+1:n}(\theta) - I_{r+1\ldots s:n}]
\end{aligned}$$

برهان: با استفاده از رابطه‌ی

$$f_{rs:n-1}(\omega;\theta) = \frac{n-s}{n} f_{rs:n}(\omega;\theta)$$

$$+\frac{s-r}{n} f_{rs+1:n}(\omega;\theta) + \frac{r}{n} f_{r+1s+1:n}(\omega;\theta)$$

که اثبات آن در دیوید [۶] می‌باشد و برهان رابطه‌ی (۳-۱) می‌توان نوشت:

$$\begin{aligned}
I_{r+1\ldots s-1|rs:n-1}(\theta) &= (n-s) I_{r+1\ldots s-1|rs:n}(\theta) \\
&+ (s-r-1) I_{r+1\ldots s|rs+1:n}(\theta) \\
&+ r I_{r+2\ldots s|r+1s+1:n}(\theta)
\end{aligned}$$

حال با استفاده از رابطه‌ی (۳-۱)، نتیجه حاصل خواهد شد.

رابطه‌ی ۳-۳ با فرض داشتن $I_{1\ldots i:n}(\theta)$ برای $i = 1, \ldots, n$، می‌توان رابطه‌ی زیر را برای محاسبه‌ی $I_{1\ldots r:d}(\theta)$ به ازای $1 \le r \le d \le n-1$ بکار برد:

$$\begin{aligned}
I_{1\ldots r:d}(\theta) &= \sum_{i=r}^{n-d+r} \frac{C_{i-1,r-1} C_{n-i-1,d-r-1}}{C_{n,d}} \\
&= \begin{cases} I_{1\ldots i:n}(\theta) & 1 \le r < d \\ \frac{r}{n} I_{1\ldots n:n}(\theta) & r = d \end{cases}
\end{aligned}$$

به طوری که $C_{n,m} = \frac{n!}{m!(n-m)!}$. رابطه‌ی فوق تعمیم رابطه‌ی (۵) برای حجم نمونه‌ی کمتر از n است.

برهان: برهان این رابطه شبیه برهان رابطه‌ی (۳-۱) است.

رابطه‌ی ۳-۴ فرض کنید برای $i = 1, \ldots, n$، دنباله‌ی $I_{1:i}(\theta)$ محاسبه شده باشد. می‌توان با استفاده از رابطه‌ی زیر، $I_{1\ldots r:n}(\theta)$ را محاسبه نمود. این رابطه یک انتگرال r گانه را به انتگرال‌های معمولی ساده می‌کند.

$$\begin{aligned}
I_{1\ldots r:n}(\theta) = \sum_{i=n-r+1}^{n} & C_{i-2,n-r-1} \qquad (9) \\
& Cn, i(-1)^{i-n+r-1} I_{1:i}(\theta), \quad 1 \le r < n
\end{aligned}$$

برهان: با استفاده از رابطه‌ی

$$\begin{aligned}
f_{s:n}(\omega;\theta) &= \sum_{i=n-s+1}^{n} C_{i-1,n-s} \\
&\times \; C_{n,i}(-1)^{i-n+s-1} f_{1:i}(\omega;\theta)
\end{aligned}$$

که اثبات آن در سریکنتن [۱۵] می‌باشد، می‌توان نوشت:

$$I_{s+1\ldots n|s:n}(\theta) = \sum_{i=n-s+1}^{n} C_{i-2,n-s-1} \times C_{n,i}(-1)^{i-n+s-1} I_{2\ldots i|1:i}(\theta)$$

و چون رابطه‌ی

$$\sum_{i=n-s+1}^{n} C_{i-2,n-s-1}C_{n-1,i-1}(-1)^{i-n+s-1} = 1$$

برقرار است، پس نتیجه حاصل خواهد شد.

نتیجه‌ی زیر از رابطه‌ی (۳-۴) بدست می‌آید:

$$I_{s\ldots n:n}(\theta) = \sum_{i=s}^{n} C_{i-2,s-2}C_{n,i}(-1)^{i-s} I_{i:i}(\theta), \quad 1 < s \le n \qquad (10)$$

رابطه‌ی ۳-۵

$$I_{rs:n}(\theta) = I_{rs:n}(\theta) + \sum_{i=s-r}^{s-1} \sum_{j=n-s+i+1}^{n} C_{i-2,s-r-2}C_{j-i-1,n-s} C_{n,j}(-1)^{n-j-r+1}[I_{1i+1:j}(\theta) - I_{1\ldots i+1:j}(\theta)]$$

برهان: با استفاده از رابطه‌ی بازگشتی زیر که اثبات آن در سریکنتن [۱۵] است، نتیجه‌ی مورد نظر بدست می‌آید:

$$f_{rs:n}(\omega;\theta) = \sum_{i=s-r}^{s-1} \sum_{j=n-s+i+1}^{n} C_{i-2,s-r-2} C_{j-i-1,n-s}Cn,j(-1)^{n-j-r+1} f_{1i+1:j}(\omega;\theta).$$

۴ اطلاع فیشر در هر مجموعه دلخواه از آماره‌های مرتب

همان‌طور که تعریف شد اطلاع فیشر برای $(X_{(r_1:n)},\ldots,X_{(r_k:n)})$ به صورت زیر نشان داده می‌شود:

$$I_{r_1\ldots r_k:n}(\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{X_{(r_{k-1}:n)}} \cdots \int_{-\infty}^{X_{(r_2:n)}} \left(\frac{\partial}{\partial\theta}\log f_{r_1\ldots r_k:n}\right)^2 dF_{r_1\ldots r_k:n}$$

در این بخش بسط ساده‌ای برای اطلاع فیشر در هر مجموعه از آماره‌های مرتب بدست می‌آوریم به طوری که این بسط شامل محاسبه‌ی انتگرال‌های حداکثر دوگانه می‌باشد.

قضیه‌ی ۴-۱ برای $1 \le r_1 < \ldots < r_k \le n$ می‌توان نوشت:

$$I_{r_1\ldots r_k:n}(\theta) = \sum_{i=1}^{k-1} I_{r_i r_{i+1}:n}(\theta) - \sum_{i=2}^{k-1} I_{r_i:n}(\theta)$$

برهان: بنابر خاصیت مارکفی آماره‌های مرتب می‌توان نوشت:

$$I_{r_i|r_1\ldots r_{i-1}:n}(\theta) = I_{r_i|r_{i-1}:n}(\theta)$$

بنابراین با تجزیه‌ی اطلاع فیشر خواهیم داشت:

$$I_{r_1\ldots r_k:n}(\theta) = I_{r_1:n}(\theta) + I_{r_2|r_1:n}(\theta) + \cdots + I_{r_k|r_{k-1}:n}(\theta)$$

که با استفاده از رابطه‌ی

$$I_{r_i|r_{i-1}:n}(\theta) = I_{r_i r_{i-1}:n}(\theta) - I_{r_i:n}(\theta)$$

رابطه‌ی (۱۰) بدست می‌آید.

قضیه‌ی فوق به صورت زیر قابل تعمیم است.

لم ۴–۱ فرض کنید $\underline{r} = \cup_{i=1}^{k} \underline{r}_i$ مجموعه‌ای از اعداد مرتب باشد به طوری که $\underline{r}_i \cap \underline{r}_j \neq \emptyset$. داریم:

$$I_{\underline{r}:n}(\theta) = \sum_{i=1}^{k-1} I_{\underline{r}_i:n}(\theta) - \sum_{i=2}^{k-1} I_{\underline{r}_i:n}(\theta)$$

به عنوان مثال می‌توان نوشت:

$$I_{12345:5}(\theta) = I_{123:5}(\theta) + I_{345:5}(\theta) - I_{3:5}(\theta)$$
$$= I_{123:5}(\theta) + I_{2345:5}(\theta) - I_{23:5}(\theta)$$
$$= \ldots$$

جدول ۱ اطلاع فیشر در مورد پارامتر مقیاسی در توزیع لجستیک را برای تمام زوج‌های آماره‌های مرتب از نمونه‌ای به حجم ۵ نشان می‌دهد به طوری که عناصر قطری برابر با $I_{i:n}(\theta)$ می‌باشند. با استفاده از این جدول می‌توانیم اطلاع فیشر در آماره‌های مرتب را برای هر مجموعه‌ی دلخواه بدست آوریم. به عنوان مثال $I_{12345:5}(\theta)$ به سادگی برابر است با:

$$۲(۳/۴۹۹۸ + ۳/۰۳۳۱) - ۲(۲/۰۹۳۲)$$
$$-۱/۷۲۹۸ = ۷/۱۴۹۶$$

که برابر است با:$5I_{1:1}(\theta)$ که دارای مقدار $۹/(\pi^2 + ۳)$ می‌باشد. به طور مشابه برای $I_{135:5}(\theta)$ می‌توان نوشت:

$$I_{135:5}(\theta) = I_{13:5}(\theta) + I_{35:5}(\theta) - I_{3:5}(\theta)$$
$$= ۴/۱۹۲۷ + ۴/۱۹۲۷ - ۱/۷۲۹۸$$
$$= ۶/۶۵۵۶$$

و به این نتیجه می‌رسیم که $(X_{(1:5)}, X_{(3:5)}, X_{(5:5)})$ دارای ۹۳/۰۹ درصد از حجم کل اطلاعات نمونه هستند.

اکنون اگر فرض کنیم $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)'$ یک بردار حقیقی مقدار p بعدی باشد می توان $\underline{I}_{r_1 \ldots r_k:n}(\theta)$ را به عنوان ماتریس اطلاع تعریف نمود به طوری که عنصر (ij) ام این ماتریس برابر است با:

$$I^{ij}_{r_1 \ldots r_k:n}(\theta) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{X_{(r_2:n)}} \left(\frac{\partial}{\partial \theta_i} \log f_{r_1 \ldots r_k:n}\right)$$
$$\left(\frac{\partial}{\partial \theta_j} \log f_{r_1 \ldots r_k:n}\right) dF_{r_1 \ldots r_k:n}$$

بنابراین قضیه‌ی (۴–۱) را می توان برای $\underline{I}_{r_1 \ldots r_k:n}(\theta)$ به شکل زیر بیان نمود.

نتیجه ۴–۱

$$\underline{I}_{r_1 \ldots r_k:n}(\theta) = \sum_{i=1}^{k-1} \underline{I}_{r_i r_{i+1}:n}(\theta) - \sum_{i=2}^{k-1} \underline{I}_{r_i:n}(\theta)$$

۵ اطلاع فیشر مجانبی

اطلاع فیشر مجانبی در چندک های نمونه در گذشته مثلاً در هلپرین [۷] و بالمر و همکارانش [۲] مورد بررسی قرار گرفته است.

تعریف: (زنگ [۱۷]) فرض کنید $(i = 1, 2, \ldots, k)$ $\frac{r_i}{n} \to p_i$ وقتی $n \to \infty$، به طوری که $0 \le p_1 < p_2 < \ldots < p_k \le 1$. اطلاع فیشر مجانبی نهفته در k چندک نمونه‌ای $(X_{r_1:n}, X_{r_2:n}, \ldots X_{r_k:n})$ در مورد پارامتر θ با $I_{p_1 p_2 \ldots p_k}(\theta)$ نشان داده شده و به شکل زیر تعریف می‌شود:

$$I_{p_1 p_2 \ldots p_k}(\theta) = \lim_{n \to \infty} \frac{1}{n} I_{r_1 r_2 \ldots r_k:n}(\theta), \qquad (۲۲)$$

که به شکل زیر نوشته می‌شود:

$$I_{p_1 p_2 \ldots p_k}(\theta) =$$
$$\sum_{i=0}^{k} \frac{\left\{\frac{\partial}{\partial \theta} F(\xi_{p_{i+1}}; \vartheta) - \frac{\partial}{\partial \vartheta} F(\xi_{p_i}; \vartheta)\right\}^2}{p_{i+1} - p_i}$$

به طوری که $p_0 = 0, p_{k+1} = 1$ و $\xi_p = F^{-1}(p;\theta)$.

۶ فضای بهینه (اپتیمال) بر اساس اطلاع فیشر

در این جا انتخاب بهینه به انتخاب بهترین زیر مجموعه از آماره‌های مرتب از نمونه‌ی داده شده گفته می‌شود به طوری که بهترین به معنای ماکزیمم نمودن اطلاع فیشر در مجموعه‌ی مورد نظر است. بنابراین انتخاب بهینه‌ی مجانبی به همین ترتیب با جایگزین نمودن اطلاع فیشر مجانبی (که در بخش قبل تعریف شد) به جای اطلاع فیشر دقیق تعریف خواهد شد.

۱.۶ حالت تک پارامتری

برای یک پارامتر می‌توان به طور ساده انتخاب بهینه را مجموعه‌ی آماره‌های مرتبی تعریف نمود که اطلاع فیشر را در بین تمام مجموعه‌های نظیر ماکزیمم کند. برای پارامتر اسکالر در توزیع لجستیک می‌توان بطور مستقیم با استفاده از جدول ۱، انتخاب بهینه به حجم ۲ را مجموعه‌ی $(X_{(1:5)}, X_{(5:5)})$ در نظر گرفت. اگر انتخاب بهینه در اندازه ۳ را در نظر بگیریم، ۱۰ انتخاب ممکن خواهیم داشت. در جدول ۲ تمام مجموعه‌های سه‌تایی و اطلاع فیشر مربوط به آنها آورده شده است. با توجه به این جدول، مجموعه‌ی $(X_{(1:5)}, X_{(3:5)}, X_{(5:5)})$ انتخاب بهینه به حجم سه از نمونه‌ی پنج تایی است.

چون اطلاع فیشر در هر زوج از آماره‌های مرتب درباره‌ی پارامتر مقیاس از توزیع نمایی و پارامتر شکل از توزیع وایبل برابر است پس انتخاب بهینه برای هر دو توزیع نیز یکسان خواهدبود. این موضوع برای اطلاع فیشر مجانبی در کولدرف [۱۰] و چنگ [۱۱] مورد بررسی قرار گرفته است.

۲.۶ حالت چند پارامتری

اگر بخواهیم $g(\theta)$ را که یک تابع اسکالر همواره از θ می‌باشد مورد بررسی قرار دهیم، می‌توانیم اطلاع فیشر حاشیه‌ای درباره‌ی پارامتر $g(\theta)$ را در $(X_{r_1:n}, X_{r_2:n}, \dots X_{r_k:n})$ که به شکل زیر تعریف می‌شود مورد بررسی قرار دهیم:

$$M_{r_1\dots r_k:n}(g(\theta)) = \frac{1}{(\nabla g)' \underline{I}^{-1}_{r_1\dots r_k:n}(\theta)(\nabla g)}$$

که در آن (∇g) گرادیان بردار $g(\theta)$ برآورد شده در θ است. بنابراین انتخاب بهینه بر اساس اطلاع فیشر حاشیه‌ای درباره‌ی $g(\theta)$ را مجموعه‌ای از آماره‌های مرتب که اطلاع فیشر حاشیه‌ای را ماکزیمم می‌نمایند تعریف می‌کنیم. اگر برای خانواده‌ی مکانی و مقیاسی $f((x-\mu)/\sigma)$ تابع $g(\mu,\sigma) = \mu + c\sigma$ را در نظر بگیریم با مسأله‌ی انتخاب بهینه برای چندک مواجهایم که انتخاب بهینه‌ی مجانبی آن توسط کیوبات و اپستاین [۹] و صالح [۱۴] مورد بررسی قرار گرفته است. در جدول ۳، انتخاب بهینه را برای تابع $g(\mu,\sigma) = \mu + c\sigma$ به ازای مقادیر مختلف c از توزیع لجستیک با پارامترهای مکان و مقیاس بررسی نموده‌ایم. از این جدول نتیجه می‌گیریم که هر چه c کوچکتر باشد انتخاب بهینه شامل آماره‌های مرتب کوچکتری شده و نسبت اطلاع مجموعه‌ی انتخاب شده افزایش می‌یابد. متذکر می‌شویم که هر انتخاب بهینه برای $\mu + c\sigma$ تقریباً نسبت به r متقارن است به طوری که

جدول ۱ — $I_{rs:n}(\theta)$ برای پارامتر مقیاس در نوریع لجستیک

r/s	۱	۲	۳	۴	۵
۱	۲/۷۶۶۶	۳/۴۹۹۸	۴/۱۹۲۷	۵/۱۸۲۹	۵/۹۹۹۸
۲		۲/۰۹۳۲	۳/۰۳۳۱	۴/۲۳۵۸	۵/۱۸۲۹
۳			۱/۷۲۹۸	۳/۰۳۳۱	۴/۱۹۲۷
۴				۲/۰۹۳۲	۳/۴۹۹۸
۵					۱/۷۶۶۶

جدول ۲ —$I_{rst:n}(\theta)$ برای پارامتر مقیاس در توزیع لجستیک

r	s	t	$I_{rst:۵}(\theta)$	r	s	t	$I_{rst:۵}(\theta)$
۱	۲	۳	۴/۴۳۹۸	۱	۴	۵	۶/۵۸۹۶
۱	۳	۴	۵/۶۴۲۹	۲	۳	۴	۴/۳۳۶۴
۱	۲	۵	۶/۵۸۹۶	۲	۳	۵	۵/۴۹۶۰
۱	۳	۴	۵/۴۹۶۰	۲	۴	۵	۵/۶۴۲۴
۱	۳	۵	۶/۶۵۵۶	۳	۴	۵	۴/۴۳۹۸

$E(X_{(r:n)}) \simeq \mu + c\sigma$.

اگر به برداری از پارامترهای نامعلوم علاقه‌مند باشیم، می‌توانیم اطلاع فیشر تعمیم یافته‌ی $\underline{I}_{r_1 \ldots r_k:n}(\theta)$ را که $|*|$ دترمینان ماتریس اطلاع است، مورد بررسی قرار دهیم. می توان اطلاع فیشر تعمیم یافته را همانند تفسیر هندسی واریانس تعمیم یافته از کار جانسون و ویچرن [۸] دنبال نمود.

انتخاب بهینه برای پارامترهای مکانی — مقیاسی توزیع لجستیک با کار چان و همکارانش [۳] برای بدست آوردن مجموعه‌ای از آماره‌های مرتب که واریانس تعمیم یافته‌ی برآوردگرهای BLU مینیمم می‌نماید، در جدول (۴) مقایسه شده است. البته در توزیع‌های متقارن اگر $\underline{r} = \{r_1, \ldots, r_k\}$ آنگاه $n + ۱ - \underline{r} = \{n + ۱ - r_1, \ldots, n + ۱ - r_k\}$ نیز انتخاب بهینه خواهد بود، که این موضوع را بالمر و همکاران [۲] و اگاوا [۱۱] مورد بررسی قرار دادند.

همان‌طور که در جدول ۴ دیده می‌شود انتخاب بهینه بر اساس ماکزیمم‌سازی اطلاع فیشر حاشیه‌ای با انتخاب چان و همکاران [۳] متفاوت بوده و دارای نسبت اطلاع بیشتری می‌باشد.

جدول ۳– انتخاب بهینه در اندازه ۴ از نمونه ای
به حجم ۲۵ از توزیع لجستیک

$g(\theta)$	انتخاب بهینه	نسبت اطلاع
μ	(۵, ۱۰, ۱۶, ۲۱)	۰/۹۶۰۷
$\mu+\sigma/۳$	(۷, ۱۳, ۱۸, ۲۳)	۰/۹۵۹۷
$\mu+۲\sigma/۳$	(۱۰, ۱۵, ۲۰, ۲۴)	۰/۹۵۴۹
$\mu+\sigma$	(۱۳, ۱۸, ۲۲, ۲۵)	۰/۹۳۹۹
$\mu+۲\sigma$	(۲, ۱۶, ۲۱, ۲۴)	۰/۹۱۱۲

جدول ۴– مقایسه انتخاب چان و همکارانش با انتخاب بهینه دراندازه ۴ از نمونه‌ای
به حجم ۲۵ از توزیع لجستیک

n	انتخاب چان و همکارانش	نسبت اطلاع	انتخاب بهینه	نسبت اطلاع
۱۶	(۱, ۴, ۱۱, ۱۵)	۰/۷۸۴۸	(۱, ۴, ۱۰, ۱۵)	۰/۷۸۶۵
۱۹	(۲, ۶, ۱۴, ۱۸)	۰/۷۷۵۹	(۲, ۶, ۱۳, ۱۸)	۰/۷۷۷۷
۲۰	(۲, ۶, ۱۴, ۱۹)	۰/۷۷۵۱	(۲, ۷, ۱۴, ۱۹)	۰/۷۷۵۹
۲۳	(۲, ۷, ۱۷, ۲۲)	۰/۷۶۷۵	(۲, ۸, ۱۷, ۲۲)	۰/۷۶۸۶
۲۴	(۲, ۷, ۱۷, ۲۳)	۰/۷۶۵۶	(۲, ۸, ۱۷, ۲۳)	۰/۷۶۵۹

مراجع

[1] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N., 1992. *A First Course in Order Statistics,* John Wiley, New York.

[2] Balmer, D.W., Boulton, M. and Sack, R.A. 1974, Optimal solution in parameter estimation problems for the cauchy distribution. *J. Amer. Statist. Assoc.,* 69, 238-242.

[3] Chan, L.K., Chan, N.N. and Cheng, S.W., 1971, Best linear unbiased estimate of the parameter of the logistic distribution based on selected order statistics. *J. Amer. Statist. Assoc.,* 66, 889-892.

[4] Cheng, S.W., 1975, A unified approach to choosing optimum quantiles for the ABLE's, *J. Amer. Statist. Assoc.,* 70, 155-159.

[5] Cole, R.H. 1951, Relation between Moment of Order Statistics, *Ann. Math. Statist.,* 22, 308-310.

[6] David, H.A., 1981, *Order Statistics,* John Wiley, New York.

[7] Helprin, M. (1952). Maximum likelihood estimation in truncated samples *Ann. Math. Statist.,* 23, 226-238.

[8] Johnson, R.A. and Wichern, D.W., 1992, Applied Multivariate Statistical Analysis, Prentice Hall, New York.

[9] Kubat, P. and Epstein, B. 1980, Estimation of quantiles of location-scale distributions based on two or three order statistics, *Technometrics* 22, 575-581.

[10] Kulldorff, G., 1973, A note on the optimum Spacing of sample quantieles from the six extreme value distribution *Ann. Statist.,* 1, 592-567.

[11] Ogawa, J., 1998, Optimal spacing of the selected sample quantiles for the joint estimation of the location and scale parameters of a symmetrc distribution, *J. Statist. Plann. and Infer.,* 70, 345-360.

[12] Park, S., 1996, Fisher information in order statistics, *J. Amer. Statist. Assoc.,* 91, 385-90.

[13] Park, S., 2005, On calculating the Fisher information in order statistics, *Statis. Papers,* 46, 293-301.

[14] Saleh, A.K., Md. E., Ali, M.M. and Umbach, D. 1983, Estimating the quantile functipon of a location-scale family of distributions based on few selected order statistics *J. Statist. Plann. and Infer.,* 8, 75-86.

[15] Srikantan ,K.S., 1962, Recurrence relation between the PDF's order statistics and some application, *Ann. Math. Statist,* 33, 169-177.

[16] Tukey, J.W., 1965, Which part of the sample contains the information? *Proceeding of National Academy of Science,* 53, 127-134.

[17] Zheng, G., 2000, On the rate of convergence of Fisher information in multiple type II censored data, *J. Japan Statist. Society,* 30, 197-204.