

محاسبه اطلاع فیشر در هر مجموعه دلخواه از آماره‌های مرتب

مهدی روزبه^۱ سید محمد 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 قرار می‌دهیم. 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 $R^1 \rightarrow \chi \times \Theta$ باشد که در آن فضای نمونه برای متغیر تصادفی X و 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 $(X_{(1:n)}, \dots, X_{(r:n)})$ به صورت زیر تعریف می‌شود:

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

^۱دانشگاه فردوسی مشهد
^۲دانشگاه فردوسی مشهد

قرار گرفتند. اما در اینجا به کمک آنها روابط بازگشتی برای اطلاع فیشدر در آماره‌های مرتب استخراج می‌کنیم که محاسبه‌ی اطلاع فیشدر در آماره‌های مرتب به کمک آنها ساده می‌شود.

هم‌چنین نتایج ساده‌ای را بدست می‌آوریم که اطلاع فیشدر در هر مجموعه از آماره‌های مرتب می‌تواند به سادگی به صورت ترکیب خطی از اطلاع فیشدر در زوج‌های آماره‌های مرتب نمایش داده شود. بنابراین با محاسبه‌ی اطلاع فیشدر در تمام مجموعه‌های دوتایی از آماره‌های مرتب به سادگی می‌توان اطلاع فیشدر 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 فیشدر در آماره‌های مرتب

چون اطلاع فیشدر دارای خاصیت جمع‌پذیری به عنوان یک اندازه از اطلاعات است رابطه‌ی (۲) به وسیله‌ی خاصیت مارکفی آماره‌های مرتب به دست می‌آید و چون $f_{r+1:n|r:n}$ 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])، $I_{r+1:n|r:n}(\theta)$ می‌تواند به صورت زیر نوشته شود:

$$I_{r+1:n|r:n}(\theta) = (n-r) \int_{-\infty}^{\infty} g(\omega; \theta) f_{r:n}(\omega; \theta) d\omega \quad (3)$$

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

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

$$I_{1:n|r:n}(\theta) = I_{1:n|r:n}(\theta) + I_{r+1:n|r:n}(\theta) \quad (2)$$

به طوری که $I_{r+1:n|r:n}(\theta)$ میانگین اطلاع شرطی در $(X_{(r+1:n)}, \dots, X_{(n:n)})$ با شرط $X_{(r:n)} = x_{(r:n)}$ می‌باشد. در معادله‌ی اخیر به محاسبه‌ی $I_{1:n|r:n}(\theta)$ علاقه‌مندیم، اما برای بدست آوردن این رابطه ابتدا باید $I_{r+1:n|r:n}(\theta)$ را مورد بررسی قرار دهیم و سپس مقدار $I_{1:n|r:n}(\theta)$ را از معادله‌ی (۲) بدست آوریم. این رویکرد را با مطالعه‌ی اطلاع فیشدر 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

به طوری که

$$I_{r+1 \dots n-1|r:n-1}(\theta) = (n-r-1) \int_{-\infty}^{\infty}$$

$$g(\omega; \theta) f_{r:n-1}(\omega; \theta) d\omega$$

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 [۶] می‌باشد، می‌توان نوشت:

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

$$\times \left\{ (n-r) \int_{-\infty}^{\infty} g(\omega; \theta) f_{r:n}(\omega; \theta) d\omega \right\}$$

$$+ \frac{r}{n} \left\{ (n-r-1) \int_{-\infty}^{\infty} g(\omega; \theta) f_{r+1:n}(\omega; \theta) d\omega \right\}$$

$$= \frac{n-r-1}{n} I_{r+1 \dots n|r:n}(\theta) + \frac{r}{n} I_{r+2 \dots n|r+1:n}(\theta)$$

بنابراین با استفاده از رابطه‌ی (۲) به نتیجه‌ی زیر

می‌رسیم:

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

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

$$I_{s \dots n-1:n-1}(\theta) = \frac{n-s}{n} I_{s \dots n:n}(\theta) + \frac{s-1}{n} I_{s+1 \dots n:n}(\theta) \quad (6)$$

اکنون با استفاده از رابطه (۲) می‌توان نوشت:

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

حال با استفاده از روابط (۷) و

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

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

بنابراین می‌توان انتگرال r گانه $I_{1 \dots r:n}(\theta)$ را بر حسبانتگرال دو گانه‌ی $n I_{1:1}(\theta) - I_{r+1 \dots n|r:n}(\theta)$ نوشت. وچون هدف ما در معادله‌ی (۲) محاسبه‌ی $I_{1 \dots r:n}(\theta)$ است، فقط کافی است $I_{r+1 \dots n|r:n}(\theta)$ را محاسبه نماییمزیرا همیشه راجع به $I_{1 \dots n:n}(\theta)$ اطلاع داریم. به عنوانمثال مقدار $I_{r+1 \dots n|r:n}(\theta)$ برای توزیع نمایی با پارامترمقیاس با استفاده از خاصیت فقدان حافظه برابر با $\frac{n-r}{\theta^2}$

است، پس می‌توان نوشت:

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

۳ روابط بازگشتی

رابطه‌ی ۱-۳

$$I_{r \dots s:n-1}(\theta) = \frac{n-s-1}{n} I_{r \dots s:n}(\theta) + \frac{s-r+1}{n} \times I_{r \dots s+1:n}(\theta) + \frac{r-1}{n} I_{r+1 \dots s+1:n}(\theta) \quad (4)$$

این رابطه، رابطه‌ی بازگشتی دیگری را هم برای اطلاع

فیشردر نخستین r آماره‌ی مرتب شامل می‌شود:

$$I_{1 \dots r:n-1}(\theta) = \quad (5)$$

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

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

که اثبات آن در دیوید [۶] می باشد و برهان رابطتهی

می توان نوشت:

(۱-۳) می توان نوشت:

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

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

رابطه‌ی ۳-۳ با فرض داشتن $I_{1 \dots i:n}(\theta)$ برای $i = 1, \dots, n$ می توان رابطتهی زیر را برای محاسبه‌ی $I_{1 \dots r:d}(\theta)$ به ازای $1 \leq r \leq d \leq n-1$ بکاربرد:

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

به طوری که $C_{n,m} = \frac{n!}{m!(n-m)!}$. رابطه‌ی فوق تعمیم رابطتهی (۵) برای حجم نمونه‌ی کمتر از n است. برهان: برهان این رابطه شبیه برهان رابطتهی (۱-۳) است.

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

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

برهان: با استفاده از رابطه‌ی

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

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

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

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

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

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

برهان: با استفاده از رابطه‌ی

$$\begin{aligned} f_{rs:n-1}(\omega; \theta) &= \frac{n-s}{n} f_{rs:n}(\omega; \theta) \\ &+ \frac{s-r}{n} f_{rs+1:n}(\omega; \theta) + \frac{r}{n} f_{r+1 s+1:n}(\omega; \theta) \end{aligned}$$

۴ اطلاع فیشر در هر مجموعه دلخواه از آماره‌های مرتب

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

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

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 $1 \leq r_1 < \dots < r_k \leq n$ می‌توان نوشت:

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

برهان: بنابر خاصیت مارکفی آماره‌های مرتب می‌توان نوشت:

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

بنابراین با تجزیه‌ی اطلاع فیشر خواهیم داشت:

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

که با استفاده از رابطه‌ی

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

رابطه‌ی (۱۰) بدست می‌آید.

قضیه‌ی فوق به صورت زیر قابل تعمیم است.

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

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

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:

$$I_{s \dots n:n}(\theta) = \sum_{i=s}^n C_{i-2, s-2} C_{n,i}(-1)^{i-s} \quad (۱۰)$$

$$I_{i:i}(\theta), \quad 1 < s \leq n$$

رابطه‌ی ۳-۵

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

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

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

عنوان ماتریس اطلاع تعریف نمود به طوری که عنصر (ij) ام این ماتریس برابر است با:

$$I_{r_1 \dots r_k:n}^{ij}(\theta) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{X(r_1:n)} \left(\frac{\partial}{\partial \theta_j} \log f_{r_1 \dots r_k:n} \right) \left(\frac{\partial}{\partial \theta_j} \log f_{r_1 \dots r_k:n} \right) dF_{r_1 \dots r_k:n}$$

بنابراین قضیه ی (۱-۴) را می توان برای $I_{r_1 \dots r_k:n}(\theta)$ به شکل زیر بیان نمود.

نتیجه ۱-۴

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

۵ اطلاع فشر مجانبی

اطلاع فشر 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 $(i = 1, 2, \dots, k)$ وقتی $n \rightarrow \infty$ به طوری که $p_1 \leq p_2 < \dots < p_k \leq 1$ اطلاع فشر مجانبی نهفته در k چندک نمونه ای $(X_{r_1:n}, X_{r_2:n}, \dots, X_{r_k:n})$ در مورد پارامتر θ با $I_{p_1 p_2 \dots p_k}(\theta)$ نشان داده شده و به شکل زیر تعریف می شود:

$$I_{p_1 p_2 \dots p_k}(\theta) = \lim_{n \rightarrow \infty} \frac{1}{n} I_{r_1 r_2 \dots r_k:n}(\theta), \quad (22)$$

که به شکل زیر نوشته می شود:

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

لم ۱-۴ فرض کنید $r = \cup_{i=1}^k r_i$ مجموعه ای از اعداد مرتب باشد به طوری که $r_i \cap r_j \neq \emptyset$ داریم:

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

به عنوان مثال می توان نوشت:

$$\begin{aligned} I_{12345:n}(\theta) &= I_{123:n}(\theta) + I_{245:n}(\theta) - I_{3:n}(\theta) \\ &= I_{123:n}(\theta) + I_{2345:n}(\theta) - I_{23:n}(\theta) \\ &= \dots \end{aligned}$$

جدول ۱ اطلاع فشر در مورد پارامتر مقیاسی در توزیع لجستیک را برای تمام زوج های آماره های مرتب از نمونه ای به حجم ۵ نشان می دهد به طوری که عناصر قطری برابر با $I_{i:n}(\theta)$ می باشند. با استفاده از این جدول می توانیم اطلاع فشر در آماره های مرتب را برای هر مجموعه ی دلخواه بدست آوریم. به عنوان مثال $I_{12345:n}(\theta)$ به سادگی برابر است با:

$$2(3/4998 + 3/0331) - 2(2/0932)$$

$$-1/7298 = 7/1496$$

که برابر است با: $I_{123:n}(\theta)$ که دارای مقدار $(3 + \pi^2)/9$ می باشد. به طور مشابه برای $I_{135:n}(\theta)$ می توان نوشت:

$$\begin{aligned} I_{135:n}(\theta) &= I_{13:n}(\theta) + I_{35:n}(\theta) - I_{5:n}(\theta) \\ &= 4/1927 + 4/1927 - 1/7298 \\ &= 6/6556 \end{aligned}$$

و به این نتیجه می رسیم که $(X_{(1:5)}, X_{(2:5)}, X_{(5:5)})$ دارای ۹۳/۰۹ درصد از حجم کل اطلاعات نمونه هستند.

اکنون اگر فرض کنیم $\theta = (\theta_1, \theta_2, \dots, \theta_p)'$ یک بردار حقیقی مقدار p بعدی باشد می توان $I_{r_1 \dots r_k:n}(\theta)$ را به

یکسان خواهد بود. 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 می باشد مورد بررسی قرار دهیم، می توانیم اطلاع فیشر حاشیه ای درباره ی پارامتر $g(\theta)$ را در $(X_{r_1:n}, X_{r_2:n}, \dots, X_{r_k:n})$ که به شکل زیر تعریف می شود مورد بررسی قرار دهیم:

$$M_{r_1, \dots, r_k; n}(g(\theta)) = \frac{1}{(\nabla g)' I_{r_1, \dots, r_k; n}^{-1}(\theta) (\nabla g)}$$

که در آن (∇g) گرادیان بردار $g(\theta)$ برآورد شده در θ است. بنابراین انتخاب بهینه بر اساس اطلاع فیشر حاشیه ای درباره ی $g(\theta)$ را مجموعه ای از آماره های مرتب که اطلاع فیشر حاشیه ای را ماکزیم می نمایند تعریف می کنیم. اگر برای خانواده ی مکانی و مقیاسی $f((x - \mu)/\sigma)$ تابع $g(\mu, \sigma) = \mu + c\sigma$ را در نظر بگیریم با مسأله ی انتخاب بهینه برای چندک مواجه ایم که انتخاب بهینه ی 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 نموده ایم. از این جدول نتیجه می گیریم که هر چه c کوچکتر باشد انتخاب بهینه شامل آماره های مرتب کوچکتری شده و نسبت اطلاع مجموعه ی انتخاب شده افزایش می یابد. متذکر می شویم که هر انتخاب بهینه برای $\mu + c\sigma$ تقریباً نسبت به r متقارن است به طوری که

به طوری که $p_0 = 0, p_{k+1} = 1$

$$\xi_p = F^{-1}(p; \theta)$$

۶ فضای بهینه (اپتیمال) بر اساس اطلاع فیشر

در این جا انتخاب بهینه به انتخاب بهترین زیر مجموعه از آماره های مرتب از نمونه ی داده شده گفته می شود به طوری که بهترین به معنای ماکزیم نمودن اطلاع فیشر در مجموعه ی مورد نظر است. بنابراین انتخاب بهینه ی 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 می توان به طور ساده انتخاب بهینه را مجموعه ی آماره های مرتبی تعریف نمود که اطلاع فیشر را در بین تمام مجموعه های نظیر ماکزیم کند. برای پارامتر اسکالر در توزیع لجستیک می توان بطور مستقیم با استفاده از جدول ۱، انتخاب بهینه به حجم ۲ را مجموعه ی $(X_{(1:5)}, X_{(5:5)})$ در نظر گرفت. اگر انتخاب بهینه در اندازه ۳ را در نظر بگیریم، ۱۰ انتخاب ممکن خواهیم داشت. در جدول ۲ تمام مجموعه های سه تایی و اطلاع فیشر مربوط به آنها آورده شده است. با توجه به این جدول، مجموعه ی $(X_{(1:5)}, X_{(3:5)}, X_{(5:5)})$ انتخاب بهینه به حجم سه از نمونه ی پنج تایی است.

چون اطلاع فیشر در هر زوج از آماره های مرتب درباره ی پارامتر مقیاس از توزیع نمایی و پارامتر شکل از توزیع وایبل برابر است پس انتخاب بهینه برای هر دو توزیع نیز

و

جدول ۱- $I_{rs:n}(\theta)$ برای پارامتر مقیاس در توزیع 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

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

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

واریانس تعمیم یافته‌ی برآوردهای BLU مینیمم

می‌نماید، در جدول (۴) مقایسه شده است. البته در توزیع‌های متقارن اگر $\underline{r} = \{r_1, \dots, r_k\}$ آنگاه $\overline{r} = \{n+1-r_1, \dots, n+1-r_k\}$ 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 ماکزیم‌سازی اطلاع فیشر حاشیه‌ای با انتخاب چان و همکاران [۳] متفاوت بوده و دارای نسبت اطلاع بیشتری می‌باشد.

اگر به برداری از پارامترهای نامعلوم علاقه‌مند باشیم، می‌توانیم اطلاع فیشر تعمیم یافته‌ی $I_{r_1 \dots r_k:n}(\theta)$ را که $|*|$ دترمینان ماتریس اطلاع است، مورد بررسی قرار دهیم. می‌توان 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

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 در اندازه ۴ از نمونه ای
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 quantiles 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 symmetric 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, *Statist. Papers*, 46, 293-301.
- [14] Saleh, A.K., Md. E., Ali, M.M. and Umbach, D. 1983, Estimating the quantile function 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.