فرهنگ و اندیشهٔ ریاضی
شماره ۴۹ (بهار ۱۳۹۱) صص. ۲۳ تا ۳۸

استنباط‌های شرطی: چرا؟ چه موقع؟ چگونه؟

علی‌رضا نعمت‌اللهی، مرجان کمالی سروستانی

چکیده

در این مقاله، به بیان اهمیت استفاده از آماره‌های فرعی و نقش آن‌ها در استنباط‌های شرطی می‌پردازیم و نشان می‌دهیم که چگونه شرطی‌کردن روی آمارهٔ فرعی می‌تواند به بهبود استنباط‌ها منجر شود. همچنین به برخی ایرادها، انتقادات و کمبودهای استنباط‌های شرطی اشاره خواهیم کرد. در ادامه، چگونگی محاسبهٔ برآوردهای مکان ــ مقیاس بهینه که با نظریهٔ شرطی ارتباط دارند، را شرح داده و توضیح می‌دهیم که چگونه شرطی‌کردن منجر به برآوردهای نقطه‌ای بهینه در بیشتر مسائل ساده و اساسی آماری می‌شود.

۱. مقدمه و تاریخچه

دو روش معمول در استنباط آماری عبارتند از این که، یا همهٔ مقادیر فضای نمونه‌ای در نظر گرفته شود و یا فقط به مقدار مشاهده شدهٔ متغیر تصادفی توجه شود. موضوع استنباط شرطی بر اساس آماره‌های فرعی[۱] برای ایجاد رابطه بین این دو روش پایه‌ریزی شده است.

ایدهٔ اساسی شرطی‌کردن دیدگاه‌های فیشر در سال ۱۹۳۴ برمی‌گردد. او معتقد است در حالت‌هایی که بعد آمارهٔ بسنده با بعد پارامتر یکسان نیست، برآوردگر درستنمایی ماکسیمم، MLE، مقداری از اطلاعات را از دست می‌دهد. از این‌رو، برای نشان دادن این‌که شرطی‌کردن می‌تواند تمامی این اطلاعات از دست‌رفته را بازیابی کند، حالت مکان ـ مقیاس را به‌عنوان نمونه مورد بررسی قرار می‌دهد و مثالی از خانوادهٔ مکان ــ متقارن (نمایی دوگانه[۲]) ارائه می‌دهد و ثابت می‌کند که

1) ancillary 2) Laplace

با شرطی‌کردن روی پیکربندی[1] (آماره‌های فرعی)، هیچ اطلاعی از بین نمی‌رود. همچنین وی توزیع شرطی MLE به‌شرط پیکربندی را در حالت کلی به‌دست می‌آورد و علاوه بر آن، بیان می‌کند در حالت‌هایی که آمارهٔ بسنده (با بعد کمتر) وجود ندارد، ممکن است بتوان با استفاده از آماره‌های فرعی، دقت برآورد را افزایش داد. تعمیم نظریهٔ فیشر در برآوردیابی نقطه‌ای، توسط پیتمن (۱۹۳۹) صورت پذیرفت. او نشان داد که چگونه می‌توان برآوردهای مکان ــ مقیاس بهینه را که با نظریهٔ شرطی ارتباط دارند، محاسبه نمود. برآوردگرهای پیتمن به‌جای مینیمم کردن خطای معمول انتگرال‌گیری شده روی کل فضای نمونه‌ای، میانگین مربع خطای شرطی را مینیمم می‌کنند. ایدهٔ شرطی‌کردن روی جنبه‌های فرعی داده‌ها، تأثیر قابل توجهی بر حل مسائل آزمون‌های معنی‌داری و برآورد فاصلهٔ اطمینان دارد. فیشر ادعا کرد که مجموعه‌های اطمینان لزوماً از تمامی اطلاعات موجود استفاده نمی‌کنند و برای اثبات این ادعا مفهوم زیرمجموعه‌های مربوطه را معرفی کرد که بعدها توسط بوهلر (۱۹۵۹) فرمول‌بندی شد. این مفهوم بدین صورت است: فرض کنید می‌خواهیم استنباطی در مورد پارامتر مجهول θ براساس متغیر تصادفی Y انجام دهیم و در این رابطه برای مجموعهٔ $C_\alpha(Y)$، ادعا می‌شود که $P_\theta(\theta \in C_\alpha(Y)) = ۱ - \alpha$. در این صورت زیرمجموعهٔ $S(Y)$ از فضای نمونه‌ای Y یک زیرمجموعهٔ مربوطه[2] است اگر برای حداقل یک $\varepsilon > ۰$ و همهٔ مقادیر پارامتر، یکی از دو رابطهٔ زیر برقرار باشد:

$$P_\theta\{\theta \in C_\alpha(Y)|S(Y)\} \geq ۱ - \alpha + \varepsilon \qquad \text{(اریبی مثبت)}$$

$$P_\theta\{\theta \in C_\alpha(Y)|S(Y)\} \leq ۱ - \alpha - \varepsilon \qquad \text{(اریبی منفی)}$$

در رابطهٔ اول، شرطی‌کردن روی زیرمجموعه‌های مربوطه باعث می‌شود که استنباط در سطحی بالاتر از مقدار تعیین‌شده باشد، بنابراین با استفاده از اطلاعات موجود در زیرمجموعه‌های مربوطه در اصل می‌توانیم عبارت شرطی دقیق‌تری بسازیم؛ به عبارت دیگر از تمامی اطلاعات موجود استفاده نکرده بودیم. مجموعهٔ

$$S(Y) = \{Y : \bar{Y}_۱^۲ + \bar{Y}_۲^۲ < n^{-۱}\Phi^{-۱}(۱ - \alpha/۲)^۲\}$$

مجموعه‌ای است با اریبی مثبت برای فاصلهٔ اطمینان پارامتر در مسألهٔ فیلر ــ کریزی[3] (که آزمونی برای نسبت دو میانگین بر اساس نمونه‌های برگرفته شده از جامعه‌های نرمال مستقل با واریانس‌های معلوم یک است)، زیرا هنگامی که داده‌ها در این مجموعه قرار می‌گیرند، فاصلهٔ اطمینان $۱۰۰(۱-\alpha)$ درصدی $(-\infty, \infty)$ در حقیقت یک فاصلهٔ ۱۰۰% است.

برعکس، شرطی‌کردن روی زیرمجموعه‌های مربوطه در رابطهٔ دوم نشان می‌دهد که سطح تعیین شده بسیار بالاست. در مسألهٔ بهرنس ــ فیشر[4] که آزمونی برای تساوی میانگین دو جامعه بر اساس نمونه‌های گرفته‌شده از جامعه‌های نرمال مستقل با واریانس‌های مجهول است، فیشر (۱۹۵۶) نشان داد که برای فاصلهٔ اطمینان تقریبی ولش (۱۹۴۷)، سطح شرطی آزمون به‌شرط

1) configuration 2) relevent 3) Feller-Crasy 4) Behrens-Fisher

$S(Y) = \{Y : S_2^2 = S_1^2\}$ از سطح کلی آزمون کمتر است و بنابراین یک زیرمجموعهٔ مربوطه با اریبی منفی است. این گونه مثال‌ها بسیاری از محققان را متقاعد کرد که انتخاب مجموعهٔ مرجع می‌تواند نتایج مهمی دربرداشته باشد. برای دیدن جزئیات بیشتر به مورگن تالر و توکی (۱۹۹۱) مراجعه کنید.

سؤالی که در اینجا می‌توان مطرح کرد در رابطه با نحوهٔ انتخاب مناسب این مجموعه‌های مرجع است. روی چه مجموعه‌هایی بایستی شرطی نمود تا به استنباط‌هایی با نتایج شرطی خوب منجر گردد؟ فیشر در بحث‌های خود ادعا می‌کند که استنباط‌ها بایستی بر اساس آماره‌های فرعی انجام شود. روش شرطی‌کردن روی آماره‌های فرعی علاوه بر این گونه ویژگی‌های خوب، دارای ویژگی‌های مهم زیر نیز هست:

۱) کاهش داده‌ها

۲) حذف پارامترهای مزاحم

۳) بازیابی اطلاعات

در ادامه، ابتدا آمارهٔ فرعی را تعریف می‌کنیم و سپس به‌اختصار این ویژگی‌ها را شرح می‌دهیم.

۲. اهمیت استفاده از آماره‌های فرعی در استنباط شرطی

فرض کنید $\theta = (\psi, \lambda)$ به‌طوری که ψ پارامتر(عدد یا بردار) مورد علاقه و λ پارامتر مزاحم می‌باشد. آمارهٔ بسندهٔ مینیمال S را که به‌صورت $S = (T, A)$ افراز شده است، در نظر بگیرید به‌طوری که

۱. توزیع A تنها بهλ بستگی داشته باشد و به ψ بستگی نداشته باشد.

۲. توزیع شرطی T به‌شرط $A = a$ برای هر a تنها به ψ بستگی داشته و به λ بستگی نداشته باشد.

آن‌گاه A را آمارهٔ فرعی برای ψ و T را بسندهٔ شرطی برای ψ، زمانی که a داده شده است، می‌نامند. یادآور می‌شویم که A باید تابعی از آمارهٔ بسندهٔ مینیمال باشد. لازم به‌ذکر است که بسیاری از نویسندگان هر تابعی از Y که توزیعش به θ بستگی ندارد را آمارهٔ فرعی می‌نامند اما در این مقاله، ما چنین آماره‌هایی را توزیع ثابت می‌نامیم. اصل شرطی (در حضور پارامتر مزاحم) بیان می‌کند استنباط دربارهٔ ψ بر اساس توزیع شرطی T به‌شرط $A = a$ صورت می‌گیرد.

همان‌گونه که در بخش ۱ بیان شد، فاصله‌های اطمینان می‌توانند خصوصیات شرطی ضعیفی داشته باشند. در این قسمت، به این سؤال پاسخ می‌دهیم که چگونه می‌توان استنباطی را انجام داد که خصوصیات شرطی خوبی داشته باشد یا به عبارت دیگر در انجام استنباط، باید روی چه کمیت‌هایی

شرطی کنیم تا مطمئن شویم که خصوصیات شرطی خوبی به‌دست می‌آید؟

۱.۲ افزایش دقت

برای افزایش دقت بهتر است روی آماره‌های فرعی شرطی کنیم، زیرا توزیع این آماره‌ها به پارامتر مجهول بستگی ندارد و بنابراین می‌تواند در دقت استنباط تأثیر بگذارد. برای روشن‌تر شدن بحث فیشر، مدل مکان بحث شده توسط پیتمن (۱۹۳۹) را در نظر بگیرید: y_1، y_2، ...، y_n یافته‌های یک نمونهٔ تصادفی دارای توزیع یکنواخت روی فاصلهٔ $(\mu - \frac{1}{2}, \mu + \frac{1}{2})$ هستند به‌طوری که μ مجهول است. قرار دهید

$$F = \{f(y\,;\mu) = \prod_{i=1}^{n} I(-\frac{1}{2} \le y_j - \mu \le \frac{1}{2}); \mu \in R\},$$

فرض کنید علاقه‌مند به انجام استنباط برای μ هستیم. تابع درستنمایی (غیرشرطی) برابر است با

$$L(\mu) = \begin{cases} 1 & y_{(n)} - \frac{1}{2} \le \mu \le y_{(1)} + \frac{1}{2} \\ 0 & \text{واگرنه} \end{cases}$$

که به‌ازای هر نقطه‌ای در بازهٔ $t \in (y_{(n)} - \frac{1}{2}, y_{(1)} + \frac{1}{2})$ ماکسیمم می‌شود. به‌طور خاص، میان‌برد $t = (y_{(1)} + y_{(n)})/2$ را به‌عنوان مقداری که تابع درستنمایی در آن ماکسیمم شده در نظر می‌گیریم. همچنین آمارهٔ $(Y_{(1)}, Y_{(n)})$ و در نتیجه $((Y_{(1)} + Y_{(n)})/2, Y_{(n)} - Y_{(1)})$ آمارهٔ بسندهٔ مینیمال می‌باشد. بنابراین طبق تعریف، $T = (Y_{(1)} + Y_{(n)})/2$ آمارهٔ بسنده شرطی و $A = Y_{(n)} - Y_{(1)}$ آمارهٔ فرعی است.

می‌توان نشان داد که تابع چگالی میان‌برد برابر است با

$$f_T(t) = n(1 - 2|t - \mu|)^{n-1}, \qquad \mu - \frac{1}{2} \le t \le \mu + \frac{1}{2},$$

$(t - u, t - l)$ یک فاصلهٔ اطمینان $100(1-\alpha)\%$ برای μ است به‌طوری که برای هر $l \le u$، رابطهٔ

$$1 - \alpha = P(l \le T - \mu \le u) = F_T(u + \mu) - F_T(l + \mu)$$

برقرار باشد که در آن، F_T تابع توزیع T است. با کمی محاسبه، فاصلهٔ اطمینان با دم‌های برابر این گونه به‌دست می‌آید:

$$\left(t - \frac{1 - \alpha^{\frac{1}{2}}}{2}, t - \frac{\alpha^{\frac{1}{2}} - 1}{2}\right)$$

به‌طور خاص، فاصلهٔ اطمینان (غیرشرطی) 100% برابر است با

$$\left(t - \frac{1}{2}, t + \frac{1}{2}\right)$$

واضح است که برد نمونه‌ای یعنی $A = Y_{(n)} - Y_{(۱)}$، حاوی اطلاعاتی است که در ساختن این فاصلهٔ اطمینان از آن استفاده نشده است. اگر برد، برابر ماکسیمم مقدار ممکن خود یعنی ۱ باشد، آن‌گاه $y_{(۱)}$ و $y_{(n)}$ دقیقاً نقاط ابتدایی و انتهایی توزیع هستند. در نتیجه μ دقیقاً نقطهٔ وسط توزیع خواهد شد. بنابراین اگر A به ۱ نزدیک باشد، تحت مدل فرض‌شده، μ به t نزدیک خواهد بود و باید یک فاصلهٔ اطمینان کوچک داشته باشیم.اگر A به صفر نزدیک باشد، μ بین $t - ۱/۲$ و $t + ۱/۲$ خواهد بود و باید فاصلهٔ اطمینان بزرگتری داشته باشیم. بنابراین به‌طور شهودی مناسب است به‌جای میانگین گرفتن از تمام مقادیر A، روی مقدار مشاهده شده A شرطی کنیم.

می‌توان نشان داد که توزیع کناری برد نمونه‌ای، یعنی A به μ بستگی ندارد و بنابراین A یک آمارهٔ فرعی است. بنابر توصیهٔ فیشر، استنباط بایستی بر مبنای توزیع شرطی $T|A$ صورت پذیرد. بر این اساس، یک فاصلهٔ اطمینان شرطی $۱۰۰(۱ - \alpha)\%$ برای پارامتر مجهول μ برابر می‌شود با

$$\left(t - (۱ - a)\frac{\alpha}{۲} - \frac{۱}{۲} + \frac{a}{۲}, t - (۱ - a)\frac{\alpha}{۲} + \frac{۱}{۲} - \frac{a}{۲}\right)$$

و تابع درستنمایی شرطی عبارت است از

$$L(\mu|a) = \begin{cases} \frac{۱}{۱-a} & t - \frac{۱}{۲} + \frac{a}{۲} \leq \mu \leq t + \frac{۱}{۲} - \frac{a}{۲} \\ ۰ & \text{واگرنه} \end{cases}$$

بنابراین همان‌گونه که مشاهده می‌شود، حالت $a = ۰$ اطلاعات کمی را در مورد μ انتقال می‌دهد و μ بین $t - ۱/۲$ و $t + ۱/۲$ قرار می‌گیرد در حالی که در حالت $a = ۱$، μ برابر با مقدار t می‌شود.

تابع درستنمایی غیرشرطی برابر است با

$$L(\mu) = \begin{cases} ۱ & t - \frac{۱}{۲} + \frac{a}{۲} \leq \mu \leq t + \frac{۱}{۲} - \frac{a}{۲} \\ ۰ & \text{واگرنه} \end{cases}$$

بنابراین تنها فاصله برای μ بر اساس تابع درستنمایی، برابر است با فاصلهٔ اطمینان ۱۰۰%:

$$(t - \frac{۱}{۲} + \frac{a}{۲}, t + \frac{۱}{۲} - \frac{a}{۲})$$

فاصلهٔ اطمینان ۱۰۰% غیرشرطی را تنها زمانی می‌توان از فاصلهٔ اطمینان فوق به‌دست آورد که $a = ۰$ باشد! بنابراین فاصلهٔ اطمینان غیرشرطی به گونه‌ای است که به نظر می‌آید همیشه در بدترین حالت هستیم.

مثلاً نمونه‌ای به اندازهٔ $n = ۲$ با $y_{(۱)} = -۰/۲۵, y_{(n)} = ۰/۴۵$ را در نظر بگیرید به‌طوری که $t = ۰/۱, a = ۰/۷$. فاصلهٔ اطمینان ۹۰ درصدی غیرشرطی برای μ برابر است با $(-۰/۲۴۲, ۰/۴۴۲)$ در حالی که فاصلهٔ شرطی ۹۰ درصدی برابر است با $(-۰/۰۳۵, ۰/۲۳۵)$ که طول‌هایشان به‌ترتیب برابر ۰/۶۸ و ۰/۲۷ است. فاصلهٔ اطمینان غیرشرطی با داده‌ها سازگار

نیست، زیرا شامل مقادیری مانند $\mu = ۰/۳$ است که برای آن $y_{(۱)} = -۰/۲۵$ قابل مشاهده نیست، زیرا

$$Y_j \sim U(۰/۳ - \frac{۱}{۲}, ۰/۳ + \frac{۱}{۲}) \qquad \Rightarrow \qquad Y_j \sim U(-۰/۲, ۰/۸)$$

ولی فاصلهٔ اطمینان شرطی، مقدار مشاهده شدهٔ آمارهٔ فرعی را در نظر می‌گیرد و چنین حالت‌های محالی را حذف می‌کند.

۲.۲. کاهش داده‌ها

یک مدل آماری برای دادهٔ Y در نظر بگیرید به‌طوری که به θ بستگی نداشته باشد و فرض کنید که $S = s(y)$ آمارهٔ بسندهٔ مینیمال برای θ باشد. قضیهٔ فاکتورگیری نشان می‌دهد که استنباط برای θ باید بر اساس چگالی S باشد. اگر بتوانیم S را به‌شکل $S = (T, A)$ بنویسیم به‌طوری که توزیع A به θ بستگی نداشته باشد، داریم

$$f_Y(y;\theta) = f_S(s;\theta) f_{Y|S}(y|s) = f_{T|A}(t|a;\theta) f_A(a) f_{Y|S}(y|s)$$

با توجه به این‌که استنباط در مورد θ بر اساس Y با استنباط در مورد θ بر اساس $T|A$ یکسان است، پس استفاده از آمارهٔ فرعی در استنباط های شرطی باعث کاهش داده‌ها می‌گردد.

$$Y \to (T, A) \to T|A$$

در صورت وجود پارامتر مزاحم λ، در بسیاری از حالت‌های مهم، چگالی داده Y را می‌توان به صورت زیر فاکتورگیری کرد:

$$f(y\,;\psi, \lambda) \propto f(t_۱|a;\psi) f(t_۲|t_۱, a;\psi, \lambda)$$

$$f(y\,;\psi, \lambda) \propto f(t_۱|t_۲, a;\psi) f(t_۲|a;\psi, \lambda)$$

که جمله‌های مستقل از پارامتر آن نادیده گرفته شده‌اند. گاهی اوقات ممکن است کمیت a وجود نداشته باشد، اما اگر موجود باشد معمولاً آن را به‌عنوان آمارهٔ فرعی در نظر می‌گیرند. اگر چنین تجزیه‌ای برقرار باشد، طبیعی است که استنباط برای ψ باید بر مبنای عبارت مقدم سمت راست باشد و آن‌گاه استنباط در مورد ψ می‌تواند بدون نیاز به برآورد λ، استنتاج شود.

۳.۲. حذف پارامتر مزاحم

گاهی اوقات می‌توان پارامتر مزاحم را با چگالی شرطی یا حاشیه‌ای که درستنمایی شرطی یا درستنمایی حاشیه‌ای نیز نامیده می شود، حذف کرد.

$$f(y\,;\psi, \lambda) = \begin{cases} f_c(y\,|a\,;\psi, \lambda) f_m(a\,;\psi) & \text{کناری} \\ f_m(a\,;\psi, \lambda) f_c(y\,|a\,;\psi) & \text{شرطی} \end{cases}$$

رابطهٔ اول در مدل‌های خانوادهٔ نمایی و رابطهٔ دوم در مدل‌های گروه تبدیلات REML، رگرسیون مقیاس و آمارهٔ فرعی پیکربندی به چشم می‌خورند و در چنین مدل‌هایی کاربرد زیادی دارند. در بسیاری از موارد، ساختن دقیقِ چنین درستنمایی‌هایی مشکل است.

۴.۲. نقش آماره‌های فرعی در بازیابی اطلاعات

فرض کنید $T_۱$ یک آمارهٔ غیربسنده برای θ و $T_۲$ آمارهٔ فرعی برای θ باشد. به عبارت دیگر، اطلاع فیشر $T_۱$ بدین گونه باشد که برای هر $\theta \in \Theta$، $I_{T_۱}(\theta) < I_Y(\theta)$ که Y تمامی داده‌هاست. آیامی‌توان تمامی اطلاعات موجود در Y را با استفاده از $T_۱$ زمانی که روی مقدار مشاهده شدهٔ $T_۲$ شرطی کنیم، بازیابی کرد؟ جواب، مثبت و روش کار، نسبتاً ساده است. این روشِ شرطی‌کردن توسط فیشر (۱۹۳۴ ـ ۱۹۵٦) در مثال معروف (Nile) مطرح شده است و بدین صورت انجام می گیرد: ابتدا چگالی شرطی $T_۱$ را در نقطهٔ $T_۱ = u$ زمانی که مقدار $T_۲ = v$ داده شده است، می‌یابیم و آن را با $g_{T_۱|v}(u;\theta)$ نشان می‌دهیم. اطلاع فیشر به‌دست آمده از این تابع چگالی شرطی، برابر است با

$$I_{T_۱|v}(\theta) = E_\theta[\{\frac{\partial}{\partial\theta}\log\{g_{T_۱|v}(T_۱;\theta)\}\}^۲]$$

در حالت کلی $I_{T_۱|v}(\theta)$ بـه ν، یعنی مقدار آمارهٔ فرعی $T_۲$ بستگی دارد. بنابراین از $I_{T_۱|T_۲}(\theta)$ حول تمامی مقادیر ممکن ν امید ریاضی می گیریم؛ به عبارت دیگر $E_{T_۲}[I_{T_۱|T_۲}(\theta)]$ را حساب می‌کنیم که با اطلاع فیشر آمارهٔ توأم $(T_۱, T_۲)$ یعنی $I_{T_۱,T_۲}(\theta) = E_{T_۲}[I_{T_۱|T_۲}(\theta)]$ برابر خواهد شد. این روش راهی برای بازگردانی اطلاعات گمشده فراهم می‌آورد. مثال زیر به این مطلب می‌پردازد:

فرض کنید $Y_۱$ و $Y_۲$ نمونهٔ تصادفی از $N(\theta, ۱)$ باشد که $\theta \in (-\infty, \infty)$ پارامتر مجهول است. می‌دانیم که $\bar{Y}$ آمارهٔ بسنده برای θ و دارای توزیع $N(\theta, \frac{۱}{۲})$ است، بنابراین $I_{\bar{Y}}(\theta) = \frac{۱}{V(\bar{Y})} = ۲$. از آنجا که $I_{Y_۱}(\theta) = ۱ < I_{\bar{Y}}(\theta)$، پس $T_۱ = Y_۱$ آمارهٔ بسنده برای θ نمی‌باشد، یعنی اگر ما تنها $Y_۱$ را در نظر بگیریم، مقداری از اطلاعات را از دست خواهیم داد. حال آمارهٔ فرعی $T_۲ = Y_۲ - Y_۱$ و توزیع توأم $(T_۱, T_۲)$ یعنی $N_۲(\theta, ۰, ۱, ۲, \rho = \frac{۱}{\sqrt{۲}})$ را در نظر بگیرید. توزیع شرطی $T_۱$ به‌شرط $T_۲ = v$ برابر با $N(\theta + \frac{۱}{۲}v, \frac{۱}{۲})$ برای $v \in (-\infty, \infty)$ است. بنابراین

$$I_{T_۱|v}(\theta) = E_{T_۱|v}[۴(T_۱ - \theta - \frac{۱}{۲}v)^۲] = ۲$$

چون این عبارت به v بستگی ندارد، $E_{T_۲}[I_{T_۱|T_۲}(\theta)] = ۲$ که با $I_{\bar{Y}}(\theta)$ برابر است. به عبارت دیگر، با شرطی‌کردن روی آمارهٔ فرعی $T_۲$، تمامی اطلاعات که برابر $I_{\bar{Y}}(\theta)$ است بازیابی می‌گردد.

به‌وضوح شرطی‌کردن روی آمارهٔ فرعی می‌تواند به استنباط‌های بهبود یافته منجر شود هرچند باسو (۱۹٦۴) نشان داد که آماره‌های فرعی یکتا نیستند و انتخاب‌های متفاوت آماره‌های فرعی می‌تواند به نتایج متفاوتی منجر شود. خوشبختانه این حالت برای روش‌های گروه تبدیلات

مکان ـ مقیاس و مدل‌های رگرسیونی اتفاق نمی‌افتد و هر پیکربندی می‌تواند مورد استفاده قرار گیرد و نهایتاً به نتایج یکسان برسد.

۳. انتقاداتی بر شرطی‌کردن روی آماره‌های فرعی

یکی از مشکلات استنباط شرطی این است که بسیاری از مدل‌های آماری، آمارهٔ فرعی ندارند یا قادر به یافتن آمارهٔ فرعی نیستیم، مثل خانواده نمایی که آمارهٔ فرعی ندارند. همچنین ممکن است در بسیاری از مدل‌ها چندین آمارهٔ فرعی وجود داشته باشد که یک نمونه از آن‌ها مدل مکان است که هر زیرمجموعه از پیکربندی، فرعی است. در ادامه، به‌اختصار به بیان این مشکلات و راه حل‌های پیشنهادی می‌پردازیم.

۱.۳. فقدان یکتایی

همان‌گونه که می‌دانیم آمارهٔ بسندهٔ مینیمال یکتا نیست اما می‌توانیم آن را به گونه‌ای تعریف کنیم که تحت تبدیلات یک به یک، یکتا باشد که این نشان می دهد آماره‌های فرعی نیز یکتا نیستند. از طرفی، ممکن است آمارهٔ فرعی $A_۱$ و $A_۲$ وجود داشته باشد، اما آمارهٔ ترکیبی $(A_۱, A_۲)$ فرعی نباشد. به زبان دیگر، توزیع حاشیه‌ای دو متغیر تصادفی، توزیع توأم آن‌ها را مشخص نمی‌کند. در این حالت، روش یکتای مشخصی برای این کار وجود ندارد. استفاده از اصل شرطی نتیجه می‌دهد که شرطی‌کردن روی یکی از آماره‌های $A_۱$ یا $A_۲$ بهتر از شرطی نکردن است، اما حداقل با دو روش موجه رو به‌رو خواهیم شد که هیچ دلیلی بر رجحان یکی بر دیگری وجود ندارد.

بدین منظور، تعریف آمارهٔ فرعی ماکسیمال، جذاب به نظر می‌رسد. علاقه‌مند به پیدا کردن آمارهٔ بسنده‌ای بودیم که تا حد امکان کوچک باشد و تمامی اطلاعات نامناسب را حذف کند. در مقابل، به دنبال آمارهٔ فرعی‌ای هستیم که تا حد امکان بزرگ باشد؛ بنابراین تمامی متغیرهای شرطی مربوط را به حساب می‌آوریم.

در بعضی حالت‌ها آمارهٔ کمکی ماکسیمال منحصر به‌فرد نیست و همین امر موجبات معرفی استنباط شرطی را پدید آورده است. هرچند در بعضی از حالت‌ها، آماره‌های فرعی ماکسیمال وجود ندارد.

۲.۳. عدم وجود

فقدان یکتایی آمارهٔ فرعی هرچند قابل ذکر است ولی در کاربردها اهمیت چندانی ندارد. مشکلات بیشتر زمانی به‌وجود می‌آیند که آمارهٔ دقیق فرعی وجود ندارد. در حالت کلی، شناسایی آمارهٔ فرعی مناسب، مهم‌تر از انتخاب بین آماره‌های فرعی متفاوت است. بدین منظور دو راه‌حل زیر ارائه شده است:

۱) سعی شود آماره‌های فرعی مجانبی ساخته شود که چندین روش برای این کار پیشنهاد شده

است. یک حالت ساده و قابل درک این است که یک مدل با چگالی $f(y;\theta,\lambda)$ ($\lambda=\lambda_{\circ}$ متناظر با مدل مورد مطالعه) در نظر گرفته شود. فرض کنید درستنمایی برای مدل بزرگتر به‌صورت تابعی از داده‌ها از طریق $(\hat{\theta},\hat{\lambda},a)$ بیان شود به‌طوری که $\hat{\theta}$, $\hat{\lambda}$ برآوردگرهای درستنمایی ماکسیمم هستند و $a_{\circ}$ یک آمارهٔ فرعی برای (θ,λ) است. $\hat{\theta}_{\circ}$ را به‌عنوان برآورد درستنمایی ماکسیمم θ هنگامی که $\lambda=\lambda_{\circ}$، در نظر بگیرید. اگر مدل مورد مطالعه صحیح باشد، آمارهٔ نسبت درستنمایی $2\{l(\hat{\theta},\hat{\lambda},a_{\circ})-l(\hat{\theta}_{\circ},\lambda_{\circ},a_{\circ})\}$ دارای توزیع مجانبی خی ـ دو می‌باشد و بنابراین به‌صورت مجانبی فرعی است.

۲) بیشتر نظریه‌های آماری جدید بر اساس استنباط‌هایی پایه‌گذاری شده‌اند که اصل شرطی بدون نیاز به تعیین دقیق و صریح آمارهٔ فرعی شرطی، مورد استفاده قرار بگیرد.

۴. برآورد نقطه‌ای از دیدگاه استنباط شرطی

در این بخش، توضیح می‌دهیم که چگونه شرطی‌کردن منجر به برآوردهای نقطه‌ای بهینه در بیشتر مسائل ساده و اساسی آماری می‌شود. یکی از مفاهیم مهم، مفهوم فرعی بودنِ پیکربندی مکان ـ مقیاس است که در حالت‌هایی که خانوادهٔ توزیع توسط گروه تبدیلات تولید شده‌اند، همیشه قادر به یافتن آماره‌های فرعی هستیم. در ابتدا به حالت مکان با مقیاس مجهول می‌پردازیم و فرمول کلی برای چگالی شرطی به‌شرط پیکربندی مکان ـ مقیاس را به‌دست می‌آوریم. این کار، شناسایی برآوردگرهای پایا را که میانگین توان دوم خطا را مینیمم می‌کنند ممکن می‌سازد. سپس حالت مقیاس با مکان مجهول را بحث می‌کنیم. بعد از آن، با کمک بحث شرطی‌کردن، برآوردهای بهینه را ارائه می‌دهیم.

۱.۴. چگالی شرطی در مدل مکان ـ مقیاس

یک ویژگی جالب استنباط شرطی روی آماره‌های فرعی مطرح شده توسط فیشر (۱۹۳۴) و پیتمن (۱۹۳۹)، جامعیت آن است به گونه‌ای که می‌توان یک فرم کلی برای چگالی شرطی و برآوردگرهای بهینه برای پارامتر مکان و مقیاس در خانوادهٔ مدل‌های مکان ـ مقیاس ارائه نمود.

لم فرض کنید Y_1، Y_2، … ، Y_n یک نمونهٔ تصادفی از خانوادهٔ $f(\frac{y-\mu}{\sigma})$ باشد. آن‌گاه شکل کلی چگالی شرطی در این خانواده به‌صورت زیر است:

$$f(t,s|c_1,\ldots,c_{n-2})=\frac{s^{n-1}\prod_{i=1}^{n}h(sc_i+st)}{\int_{-\infty}^{\infty}\int_{\circ}^{\infty}s^{n-1}\prod_{i=1}^{n}h(sc_i+st)\,ds\,dt}.$$

اثبات. برای خانوادهٔ مدل‌های مکان ـ مقیاس

$$F=\{f(y\,;\mu,\sigma)=\prod_{i=1}^{n}\frac{1}{\sigma}h(\frac{y_i-\mu}{\sigma}),-\infty<y_i<\infty,\mu\in R,\sigma>\circ\}. \qquad (۱.۴)$$

برآوردگرهای پایای $(\hat{\mu},\hat{\sigma})$ را که دو خاصیت زیر برای آن‌ها برقرار است، در نظر بگیرید:

$$\hat{\mu}(ay+b)=a\hat{\mu}(y)+b \qquad , \qquad \hat{\sigma}(ay+b)=a\hat{\sigma}(y). \tag{۲.۴}$$

بنابراین، پیکربندی مکان ــ مقیاس $\{C_i=(Y_i-\hat{\mu})/\hat{\sigma}, i=۱,\ldots,n\}$ در روابط زیر صدق می‌کند:

$$\hat{\mu}(C)=۰ \quad , \quad \hat{\sigma}(C)=۱$$

و در نتیجه $C_{n-۱}$ و C_n را می‌توان برحسب $C_۱$، ...، $C_{n-۲}$ بیان نمود، یعنی فضای تمام پیکربندی‌ها دقیقاً $(n-۲)$ ــ بعدی است. همچنین از رابطهٔ (۲.۴) برای $a>۰$ به‌دست می‌آید که

$$C_i(aY+b)=\frac{aY_i+b-\hat{\mu}(aY+b)}{\hat{\sigma}(aY+b)}=\frac{Y_i-\hat{\mu}(Y)}{\hat{\sigma}(Y)}=C_i(Y)$$

یعنی پیکربندی نسبت به تبدیلات خطی پایا است و دارای توزیعی است که به (μ,σ) بستگی ندارد. به عبارت دیگر، پیکربندی، یک آمارهٔ فرعی است و مطابق با توصیهٔ فیشر باید استنباط‌مان را بر اساس توزیع شرطی $(\hat{\mu},\hat{\sigma})$ به‌شرط پیکربندی انجام دهیم.

بر این اساس، چگالی توأم $(\hat{\mu},\hat{\sigma},C_۱,\ldots,C_{n-۲})$ برابر است با

$$f(\hat{\mu},\hat{\sigma},c_۱,\ldots,c_{n-۲})=k(c,n)\frac{\hat{\sigma}^{n-۲}}{\hat{\sigma}}\prod_{i=۱}^{n}h(\frac{\hat{\sigma}c_i+\hat{\mu}-\mu}{\sigma}).$$

حال تبدیلات زیر را در نظر بگیرید:

$$\hat{\mu}=\sigma st+\mu, \quad \hat{\sigma}=\sigma s, \quad c_i=c_i$$

که $(\hat{\mu},\hat{\sigma},C_۱,\ldots,C_{n-۲})$ را به $(T,S,C_۱,\ldots,C_{n-۲})$ می‌نگارد. توجه کنید که $S=\frac{\hat{\sigma}}{\sigma}$ & $T=\frac{\hat{\mu}-\mu}{\hat{\sigma}}$ کمیت‌های محوری هستند. ژاکوبین این تبدیلات عبارت است از $s\sigma^۲$. بنابراین چگالی توأم $(T,S,C_۱,\ldots,C_{n-۲})$ برابر است با

$$f(t,s,c_۱,\ldots,c_{n-۲})=k(c,n)s^{n-۱}\prod_{i=۱}^{n}h(sc_i+st)$$

اگر تعریف کنیم $\hat{\mu}=\sigma r+\mu$، $\hat{\sigma}=\sigma s$ و $c_i=c_i$، آن‌گاه ژاکوبین برابر با $\sigma^۲$ می‌شود و چگالی توأم به‌شکل

$$f(r,s,c_۱,\ldots,c_{n-۲})=k(c,n)s^{n-۲}\prod_{i=۱}^{n}h(sc_i+r)$$

خواهد بود. این دو تابع چگالی توأم در دو دستگاه مختصات زیر هستند:

$$y_i=sc_i+r$$
$$y_i=s(c_i+t), \quad r=st$$

همان‌گونه در ادامه خواهیم دید، هر یک از این دو دستگاه برحسب مورد، به سادگی محاسبات کمک می‌کنند. بنابراین چگالی شرطی (S,T) به‌شرط پیکربندی، عبارت است از

$$f(t,s|c_1,\ldots,c_{n-2}) = \frac{s^{n-1}\prod_{i=1}^{n} h(sc_i+st)}{\int_{-\infty}^{\infty}\int_{0}^{\infty} s^{n-1}\prod_{i=1}^{n} h(sc_i+st)\,ds\,dt} \tag{۳.۴}$$

و تابع توزیع شرطی محوری $S=\frac{\hat{\sigma}}{\sigma}\ \ \&\ \ T=\frac{\hat{\mu}-\mu}{\hat{\sigma}}$ برابر است با

$$K(t_0,s_0|c_1,\ldots,c_{n-2}) = \frac{\int_{\hat{\mu}-\hat{\sigma}t_0}^{\infty}\int_{\hat{\sigma}s_0}^{\infty}\frac{1}{v^{n+1}}\prod_{i=1}^{n} h(\frac{y_i-t}{v})\,dv\,du}{\int_{-\infty}^{\infty}\int_{0}^{\infty}\frac{1}{v^{n+1}}\prod_{i=1}^{n} h(\frac{y_i-t}{v})\,dv\,du} = 1-G(\hat{\mu}-\hat{\sigma}t_0,\frac{\hat{\sigma}}{s_0};y) \tag{۴.۴}$$

که در آن

$$G(t,s;y) = 1-\frac{\int_{t}^{\infty}\int_{s}^{\infty}\frac{1}{v^{n+1}}\prod_{i=1}^{n} h(\frac{y_i-t}{v})\,dv\,du}{\int_{-\infty}^{\infty}\int_{0}^{\infty}\frac{1}{v^{n+1}}\prod_{i=1}^{n} h(\frac{y_i-t}{v})\,dv\,du} \tag{۵.۴}$$

اگرچه تابع توزیع شرطی توأم (۴.۴) به انتخاب $(\hat{\mu},\hat{\sigma})$ (و بنابراین به پیکربندی) بستگی دارد، اما فاصله‌های اطمینان شرطی به انتخاب $(\hat{\mu},\hat{\sigma})$ بستگی ندارند. یک فاصلهٔ اطمینان شرطی $100(1-\alpha)\%$ برای μ در (۱.۴) عبارت است از

$$[\hat{\mu}-\hat{\sigma}K^{-1}(q_2,\infty;c),\hat{\mu}-\hat{\sigma}K^{-1}(q_1,\infty;c)] = [G^{-1}(1-q_2,\infty;y),G^{-1}(1-q_1,\infty;y)]$$

این فاصلهٔ اطمینان نسبت به انتخاب $(\hat{\mu},\hat{\sigma})$ (و در نتیجه انتخاب پیکربندی) پایا است، بنابراین برای هر جفت از برآوردگرهای پایای $(\hat{\mu},\hat{\sigma})$ یکسان خواهد بود، زیرا G تعریف شده در (۵.۴) نسبت به انتخاب $(\hat{\mu},\hat{\sigma})$ پایا می‌باشد.

۲.۴. برآوردگرهای مکان پیتمن

لم در مدل مکان با مقیاس مجهول، تحت تابع زیان درجه دو، برآوردگر پایا با کمترین میانگین توان دوم خطا:

الف) در دستگاه $y_i=r+sc_i$ عبارت است از

$$T_F(c)=\frac{-E_F(rs|c)}{E_F(s^2|c)}.$$

ب) در دستگاه $y_i=s(c_i+t)$ عبارت است از

$$T_F(c)=\frac{-E_F(ts^2|c)}{E_F(s^2|c)}.$$

اثبات. اگر T برآوردگر پایای مکان باشد، همان‌گونه که دیدیم، به‌وسیلهٔ پیکربندی $\{C_i = (Y_i - \hat{\mu})/\hat{\sigma}, i = ۱, \ldots, n\}$ بایستی در رابطهٔ زیر صدق کند:

$$T(y_۱, \ldots, y_n) = \hat{\mu}(y_۱, \ldots, y_n) + \hat{\sigma}(y_۱, \ldots, y_n)T(c_۱, \ldots, c_n)$$

یعنی اطلاع از $T(c_۱, \ldots, c_n)$ برای محاسبهٔ $T(y_۱, \ldots, y_n)$ کافی است. بنابراین مسأله به تعیین $T(c_۱, \ldots, c_n)$ برای تمام $c = (c_۱, \ldots, c_n)$ تبدیل می‌شود به‌طوری که میانگین توان دوم خطای برآوردگر به‌دست آمده $T(y_۱, \ldots, y_n)$ مینیمم باشد. این نکته نقش تعیین کنندهٔ استفاده از خاصیت پایایی را در استنباط‌های شرطی بیان می‌کند. لازم است که یک دستگاه مختصات برای تمام نمونه‌هایی که پیکربندی یکسانی دارند، معرفی کنیم و چنان‌که در بخش قبلی دیدیم، می‌تواند به دو روش ساده زیر به‌دست آید:

برای r حقیقی و s حقیقی مثبت، داشته باشیم

$$y_i = r + sc_i \quad ; \ (i = ۱, \ldots, n)$$

یا برای t حقیقی و s حقیقی مثبت، داشته باشیم

$$y_i = s(c_i + t) \ ; \ (i = ۱, \ldots, n)$$

این دو انتخاب به‌وسیلهٔ تبدیل $r = st$ با هم مرتبط هستند. میانگین توان دوم خطا تحت نمونه‌گیری از توزیع $F(y)$ برابر است با

$$MSE_F(T) = E_F((T(y_۱, \ldots, y_n) - \mu(F))^۲)$$

که $\mu(F)$ مرکزتقارن $F(y)$ یا هدف از پیش مشخص شدهٔ دیگری است. با شرطی‌کردن روی پیکربندی C، داریم

$$MSE_F(T) = E_{F^*}(E_F((T(y_۱, \ldots, y_n) - \mu(F))^۲|c_۱, \ldots, c_n))$$

که در آن $F^*(c)$ توزیع روی فضای پیکربندی‌های استنتاج شده توسط نمونه‌گیری از $F(y)$ است. این فرمول، محاسبهٔ میانگین مربع خطا را به دو قسمت تقسیم می‌کند. امید ریاضی داخلی، یک امید شرطی است و امید ریاضی خارجی، وابسته به توزیع حاشیه‌ای F^* است. این عبارت را می‌توان به‌صورت ساده‌تر زیر نوشت:

$$MSE_F(T) = E_{F^*}(cMSE_F(T(y_۱, \ldots, y_n)|c_۱, \ldots, c_n)$$

که در آن $cMSE_F(.|c)$ نشان دهندهٔ میانگین توان دوم خطای شرطی برآوردگر به‌شرط پیکربندی c است. از آنجا که تمام این کمیت‌ها، مثبت هستند و می‌توانیم برای هر $c = (c_۱, \ldots, c_n)$ به‌طور

جداگانه $T(c)$ انتخاب کنیم، برآوردگر T_F که $MSE_F(T)$ را مینیمم می‌کند می‌تواند این کار را فقط با مینیمم کردن همزمان تمام میانگین توان دوم خطاهای شرطی انجام دهد.

اگر شکل مرجع $F(y)$ را به‌گونه‌ای انتخاب کنیم که حول صفر متقارن باشد یا در حالت کلی‌تر $\mu(F) = \text{۰}$، آن‌گاه فرمول میانگین توان دوم خطای T به‌شکل زیر ساده می‌شود:

$$cMSE_F(T|c) = E_F(r^{\text{۲}}|c) + \text{۲}T(c)E_F(rs|c) + (T(c))^{\text{۲}}E_F(s^{\text{۲}}|c)$$

اندیس F نشان می‌دهد که امید شرطی، تحت این فرض که نمونه از F می‌باشد، محاسبه شده است. برای پیکربندی c، $cMSE_F(T|c)$ توسط کمیت

$$T_F(c) = \frac{-E_F(rs|c)}{E_F(s^{\text{۲}}|c)}$$

مینیمم می‌شود که با جانشینی $T_F(c)$ خواهیم داشت

$$cMSE_F(T_F|c) = E_F(r^{\text{۲}}|c) + T_F(c)E_F(rs|c).$$

اثبات قسمت (ب) مانند قبل است. در این حالت، برآوردگر بهینه برابر می‌شود با

$$T_F(c) = \frac{-E_F(ts^{\text{۲}}|c)}{E_F(s^{\text{۲}}|c)}.$$

این برآوردگرهای پایا با کمترین میانگین توان دوم خطا(MRE) را برآوردگرهای پیتمن می‌نامند، زیرا برای اولین بار پیتمن آن‌ها را با استفاده از تکنیک استنباط شرطی به‌دست آورد. روش دیگر و متداول‌تر برای محاسبهٔ برآوردگر پیتمن را می‌توان در فصل سوم از لهمن (۱۹۸۳) یافت.

در عمل، یافتن فرم‌های بسته برای برآوردگرهای T_F پیتمن تقریباً مشکل است. فقط در موارد محدودی، می‌توان حاصل انتگرال دوگانه را به‌صورت تحلیلی به‌دست آورد و در اکثر موارد باید از روش‌های عددی استفاده کرد.

۳.۴. برآوردگرهای مقیاس پیتمن

در برخی موارد، در برآورد پارامتر مقیاس بهتر است تابع زیان را تغییر دهیم. توزیع نمونه‌ای برآوردگرهای مقیاس معمولاً چوله به راست هستند و میانگین مربع خطا، لزوماً تابع زیان مناسبی نیست.

تعریف برآوردگر $S(y_{\text{۱}},\ldots,y_n)$ برای σ بهینه است اگر $E_F((\log S - \tau(F))^{\text{۲}})$ را مینیمم کند هنگامی که $\tau(F) = \log\sigma(F)$ یک هدف از پیش مشخص شده باشد (به یاد داشته باشید که $\sigma(F)$ لزوماً گشتاور دوم حول میانگین نیست).

لم تحت گروه مقیاس در خانوادهٔ توزیع‌های مقیاس، برآوردگر بهینهٔ پیتمن برای پارامتر مقیاس عبارت است از

$$S_F(c_1,\dots,c_n) = \exp E_F(\log \frac{\sigma(F)}{s}|c) = \exp(\int_0^\infty -\log(s)k_F(s,r|c)\,ds).$$

اثبات. مایل به پیداکردن برآوردگری بهینه مانند برآوردگر S در میان برآوردگرهای مکان پایا و مقیاس پایا هستیم. با شرطی‌کردن روی پیکربندی $(c_1,\dots,c_n)$، باید داشته باشیم $S(r+sc_1,\dots,r+sc_n) = sS(c_1,\dots,c_n)$. میانگین توان دوم خطای برآوردگر S برابراست با

$$\begin{aligned} MSE_F(S) &= E_{F^*}(E_F((\log S(y_1,\dots,y_n) - \tau(F))^2|c)) \\ &= E_{F^*}(E_F((\log s - \tau(F) + \log S(c_1,\dots,c_n))^2|c)). \end{aligned}$$

انتخاب بهینهٔ S_F از $S(c_1,\dots,c_n)$ کمیتی است که عبارت زیر را مینیمم کند:

$$E_F((\log s - \tau(F) + \log S(c_1,\dots,c_n))^2|c)$$

و این اتفاق وقتی رخ می‌دهد که

$$\log S(c_1,\dots,c_n) = \tau(F) - E_F(\log(s)|c) = E_F(\log \tfrac{\sigma(F)}{s}|c) = -E_F(\log \tfrac{s}{\sigma(F)}|c)$$

اگر فرض کنیم $F(y)$ استاندارد شده است، یعنی در دو شرط زیر صدق می‌کند:

$$\mu(F) = 0, \sigma(F) = 1$$

آن‌گاه برآوردگر پیتمن برای پارامتر مقیاس σ عبارت است از

$$S_F(c_1,\dots,c_n) = \exp E_F(\log \frac{\sigma(F)}{s}|c) = \exp\Big(\int_0^\infty -\log(s)k_F(s,r|c)\,ds\Big).$$

۴.۴. موارد خاص (دارای فرم بسته)

در معرفی برآوردگرهای مکان پیتمن، دیدیم که امیدهای شرطی مورد نیاز برای تعیین برآوردهای بهینه برای پیکربندی c، یعنی $T_F(c)$ و میانگین توان دوم خطای این برآوردها عبارتند از

$$E_F(t^2s^2|c), E_F(rs|c) = E_F(ts^2|c), E_F(s^2|c)$$

در وضعیت‌های بسیاری، این سه امید شرطی باید با استفاده از روش های عددی تقریب زده شوند. هرچند برای موارد محدودی، انتگرال‌ها به‌صورت تحلیلی و به فرم بسته به‌دست می‌آیند. در این بخش، مثالی از حالتی که در آن، مقدار انتگرال را می‌توان به‌صورت تحلیلی محاسبه گردد، ارائه می‌گردد.

مثال (**حالت گوسی**) در حالت گوسی، چگالی شرطی متناسب است با

$$s^{n-1}\exp(\frac{-s^{2}}{2}(\sum_{i=1}^{n}(c_i-\bar{c})^{2}+n(t+\bar{c})^{2}))$$

به‌طوری که $\bar{c}=\frac{1}{n}\sum_{i=1}^{n}c_i$. مقدار ثابت نرمال این چگالی شرطی برابر است با

$$k=(\frac{2\pi}{n})^{\frac{1}{2}}\int_{0}^{\infty}s^{n-2}\exp(\frac{-s^{2}}{2}(\sum_{i=1}^{n}(c_i-\bar{c})^{2})ds.$$

برای تعیین $T_F(c)=\frac{-E_F(ts^{2}|C)}{E_F(s^{2}|C)}$ با محاسباتی ساده، داریم

$$E_F(s^{2}|c)=\frac{(n-1)}{\sum_{i=1}^{n}(c_i-\bar{c})^{2}}$$

و

$$E_F(rs|c)=E_F(ts^{2}|c)=(-\bar{c})E_F(s^{2}|)=(-\bar{c})(n-1)/\sum_{i=1}^{n}(c_i-\bar{c})^{2}$$

که در نتیجه $T_F(c)$ برابر می‌شود با

$$T_F(c)=\frac{-E_F(ts^{2}|c)}{E_F(s^{2}|c)}=\bar{c}$$

به‌طوری که $F(x)$ تابع توزیع نرمال است. برای محاسبهٔ $cMSE_F(T_F|c)$ نیاز به محاسبهٔ $E_F(t^{2}s^{2}|c)$ داریم که برابر است با

$$E_F(r^{2}|c)=E_F(t^{2}s^{2}|c)=\frac{1}{n}+(\bar{c})^{2}E_F(s^{2}|c)=\frac{1}{n}+\frac{(\bar{c})^{2}(n-1)}{\sum_{i=1}^{n}(c_i-\bar{c})^{2}}$$

بنابراین

$$cMSE_F(T_F|c)=E_F(ts^{2}|c)T_F(c)+E_F(t^{2}s^{2}|c)=\frac{1}{n}$$

که این مقدار برای هر پیکربندی c یکسان است.

این مثال یک ساختار ساده را نمایش می دهد. در این حالت توانستیم پیکربندی را به گونه‌ای انتخاب کنیم که در روابط زیر صدق کند:

$$c_1+\cdots+c_n=0 \quad \& \quad c_1^{2}+\cdots+c_n^{2}=n-1$$

برای این انتخاب، همهٔ پیکربندی‌ها دقیقاً یکسان و مشابه رفتار می‌کنند. بهینگی گوسی پیشنهاد می‌کند که از میانگین حسابی استفاده کنیم. علاوه بر آن، دقت ما به پیکربندی c بستگی ندارد. در این صورت انتگرال بند قبل، به‌آسانی قابل محاسبه است و خواهیم داشت $MSE_F(T_F)=\frac{1}{n}$.

مراجع

[1] Basu, D., "Recovery of Ancillary Information", *Sankhya*, **26**(1984), 3-16.

[2] Buehler, R. J., "Some Validity Criteria for Statistical inference", *Ann. Math. Statist.*, **30**(1959), 845-863.

[3] Fisher, R. F., "Two New Properties of Mathematical Likelihood", *Proc. Roy. Soc. A*, **144**(1934) , 285-307

[4] Casella, G. and Berger, R. L., *Statistical Inference*, Duxbury, Pasific Grove, CA, 2002.

[5] Davison, A. C., *Likelihood Theory*, Brisbane, 2007.

[6] Fisher, R. F., "On a Test of Significance in Pearson's Biometrika Tables", *J. Roy. Statist. Soc. B*, **18**(1939), 56-60.

[7] Fraser, D. A. S., "Ancillaries and Conditional Inference", *Statistical Science*, **19**(2004) , No. 2, 333-369.

[8] Lehmanne, E. L., *Testing Statistical Hypotheses*, Chapman and Hall, New York, Londonm 1991.

[9] Lehmanne, E. L., *Theory of Piont Estimation*, John Wiley, New York, 1983.

[10] Morgenthaler, S. and Tukey, J. W., *Configural Polysampling: A Route to Practical Robustness*, John Wiley, New York, 1991.

[11] Pitman, E. J. G., "The Estimation of the Location and Scale Parameters of a Continuous Population of any Given Form", Biometrika, **30**(1939), 391-421.

[12] Welsh, A. H., *Aspect of Statistical Inference*, John Wiley, New York, 1996.

[13] Young, G. A. and Smith, R. L., *Essentials of Statistical Inference*, Cambridge University Press, New York, 2005.

علی‌رضا نعمت‌اللهی، nematollahi@susc.ac.ir
مرجان کمالی سروستانی، mks_82000@yahoo.com
دانشگاه شیراز، بخش آمار