بسم الله الرحمن الرحیم

کلیه حقوق مادی مترتب بر نتایج مطالعات، ابتکارات و
نوآوری‌های ناشی از تحقیق موضوع این پایان‌نامه
متعلق به دانشگاه رازی است.

دانشگاه رازی

دانشکده علوم پایه

گروه آمار

پایان‌نامه جهت اخذ درجه کارشناسی ارشد

رشته آمار ریاضی

عنوان :

طرح‌های D - بهینه برای مدل رگرسیون پواسن چند گانه با ضرایب تصادفی

استاد راهنما:

دکتر مهرداد نیاپرست

نگارش:

داریوش نادری

شهریور ۱۳۹۲

دانشگاه رازی
دانشکده علوم پایه
گروه آمار

پایان‌نامه جهت اخذ درجه کارشناسی ارشد
رشته آمار

نام دانشجو:
داریوش نادری

تحت عنوان :

طرح‌های D - بهینه برای مدل رگرسیون پواسن چند گانه با ضرایب تصادفی

در تاریخ ۹۲/۰۶/۲۶ توسط هیأت داوران زیر بررسی و با درجه **عالی** به تصویب نهایی رسید.

۱ - استاد راهنمای پایان‌نامه	دکتر مهرداد نیاپرست	با مرتبه‌ی علمی استادیار	امضاء:
۲ - استاد داور داخل گروه	دکتر حبیب جعفری	با مرتبه‌ی علمی استادیار	امضاء:
۳ - استاد داور داخل گروه	دکتر محمد مرادی	با مرتبه‌ی علمی استادیار	امضاء:

خدایا...

به من زیستنی عطا کن که در لحظه مرگ، بر بی‌ثمری لحظه‌ای که برای زیستن گذشته است، حسرت نخورم و مُردنی عطا کن که بر بیهودگیش، سوگوار نباشم. بگذار تا آن را، خود انتخاب کنم، اما آنچنان که تو دوست می‌داری.

تو می‌دانی و همه می‌دانند که شکنجه دیدن بخاطر تو، زندانی کشیدن بخاطر تو و رنج بردن به پای تو تنها لذت بزرگ زندگی من است، از شادی توست که من در دل می‌خندم، از امید رهایی توست که برق امید در چشمان خسته‌ام می‌درخشد و از خوشبختی توست که هوای پاک سعادت را در ریه‌هایم احساس می‌کنم. نمی‌توانم خوب حرف بزنم. نیروی شگفتی را که در زیر کلمات ساده و جمله‌های ضعیف و افتاده، پنهان کرده‌ام دریاب، دریاب.

تو می‌دانی و همه می‌دانند که زندگی از تحمیل لبخندی بر لبان من، از آوردن برق امیدی در نگاه من، از برانگیختن موج شعفی در دل من، عاجز است.

تو، چگونه زیستن را به من بیاموز، چگونه مردن را خود خواهم آموخت.

به من توفیق تلاش در شکست، صبر در نومیدی، رفتن بی‌همراه، جهاد بی‌سلاح، کار بی‌پاداش، فداکاری در سکوت، دین بی‌دنیا، مذهب بی‌عوام، عظمت بی‌نام، خدمت بی‌نان، ایمان بی‌ریا، خوبی بی‌نمود، گستاخی بی‌خامی، قناعت بی‌غرور، عشق بی‌هوس، تنهایی در انبوه جمعیت، و دوست داشتن بی‌آنکه دوست بداند، روزی کن.

اگر تنهاترین تنها شوم، باز خدا هست

او جانشین همه نداشتن‌هاست...

سپاس‌گزاری...

سپاس خداوندگار حکیم را که با لطف بی‌کران خود، آدمی را زیور عقل آراست.

در آغاز وظیفه خود می‌دانم از زحمات بی‌دریغ اساتید راهنمای خود، جناب آقایان دکتر مهرداد نیاپرست ، صمیمانه تشکر و قدردانی کنم که قطعاً بدون راهنمایی‌های ارزنده ایشان، این مجموعه به انجام نمی‌رسید و از دکتر ابوالقاسمی و دکتر نیامرادی که در طول دوران تحصیل بنده را یاری نمودند، تشکر می‌کنم.

بر خود لازم می‌دانم که از دوستان عزیزم آقایان: علی‌نقی رضایی، صابر بالایی و عباس همتی به پاس عاطفه سرشار و گرمای امیدبخش وجودشان، تشکر کنم.

در پایان، بوسه می‌زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا، ستایش می‌کنم وجود مقدس برادر و خواهر عزیزم را که در این سردترین روزگاران، بهترین پشتیبان من بودند.

داریوش نادری

شهریور ۱۳۹۲

تقدیم به

برادر و خواهر عزیزم

چکیده

یکی از مباحثی که اخیراً در مسائل کاربردی مورد توجه بعضی از آماردانان قرار گرفته شده است، استفاده از طرح‌های آزمایش بهینه برای انجام آزمایش‌های آمار است. این طرح‌ها بر اساس ماکسیمم کردن توابعی از ماتریس اطلاع فیشر برای پارامتر به‌دست می‌آیند. از آن‌جایی‌که در مدل‌های آمیخته خطی تعمیم‌یافته، نمی‌توان فرم صریحی برای تابع درستنمایی و در نتیجه ماتریس اطلاع فیشر به‌دست آورد، بنابراین از روش‌های تقریبی بجای روش درستنمایی استفاده می‌شود. در اینجا از روش شبه‌درستنمایی جهت به‌دست آوردن تقریبی از ماتریس اطلاع برای مدل‌های رگرسیون پواسن با ضرایب تصادفی استفاده شده است. این مدل‌ها، حالت‌های خاصی از مدل‌های آمیخته خطی تعمیم‌یافته هستند.

در این پایان نامه طرح‌های D-بهینه را برای رگرسیون پواسن در حالتی که اثرات ثابت و تصادفی باشند مورد بحث قرار داده می‌شود و در پایان مدل‌های خاصی از رگرسیون پواسن دوگانه با ضرایب تصادفی را بررسی خواهد شد و استراتژیی برای بدست آوردن طرح های بهینه مدل های پواسن چند گانه با اثرات تصادفی ارائه می دهیم.

کلمات کلیدی:

اثر تصادفی، روش درستنمایی، روش شبه‌درستنمایی، طرح آزمایش بهینه، مدل رگرسیون پواسن، D- بهینه.

فهرست مطالب

عنوان | صفحه

لیست جداول

لیست تصاویر

پیشگفتار

در انجام یک آزمایش، مسائلی چون هزینه و میزان اعتماد به نتیجه آزمایش، پژوهشگران را به این سمت سوق داد که قبل از انجام آزمایش، طرح آزمایش را برای رسیدن به بهترین پاسخ به‌دست آورند.

طرح‌های بهینه، کلاسی از طرح آزمایشها هستند که نسبت به برخی معیارهای آماری بهینه هستند. اولین بار اسمیت(۱۹۱۸)طرح بهینه را برای یک مدل خطی چندجمله‌ای ارائه کرد.

بهینگی یک طرح، در مدل آماری به داده‌ها وابسته است. برای مدل‌های خطی به‌دلیل سادگی، پژوهش‌های زیادی صورت گرفته، که در کتاب‌های سیلوی (۱۹۸۰)،فدرو(۱۹۷۲) و پوکلسهایم(۱۹۴۸) به‌طور جامع بررسی شده است. اما از نظر کاربردی مدل‌های خطی تعمیم‌یافته و غیر خطی از اهمیت بیشتری برخوردارند. چرنوف(۱۹۵۳)طرح بهینه موضعی را برای این مدل‌ها پیشنهاد داد.

اخیراً طرح بهینه برای مدل‌هایی که در کنار اثرات ثابت دارای اثر تصادفی نیز هستند مورد توجه قرار گرفته است. برای مدل‌های آمیخته خطی نتایجی در فدرو و هکل(۱۹۹۷) و اشملتر(۲۰۰۷) بیان شده است.

در مدل‌های آمیخته خطی تعمیم‌یافته، طرح بهینه از روش‌های معمول قابل محاسبه نیست. زیرا در این مدل‌ها ماتریس اطلاع شکل صریحی ندارد. خطی کردن و برخی روش‌های دیگر برای حل این مشکل بیان شده، اما این روش‌ها قابل اعتماد نیستند. نیاپرست(۲۰۰۹)روش شبه‌درستنمایی را برای حالت خاصی از این مدل ارائه کرد و طرح D-بهینه را برای مدل رگرسیون پواسن با عرض از مبدأ تصادفی به‌دست آورد. و همچنین نیاپرست و شوابه (۲۰۱۳) فرضیه محدب بودن ماتریس شبه اطلاع را برای مدل رگرسیون پواسن با ضرایب تصادفی ارائه دادند.

در فصل اول این پایان‌نامه مرور مختصری بر مدل‌های رگرسیونی شامل مدل‌های خطی، مدل‌های خطی تعمیم‌یافته، مدل‌های آمیخته خطی ،مدل‌های آمیخته خطی تعمیم‌یافته خواهیم داشت و در آخر ساختار مدل رگرسیون پواسن با ضرایب تصادفی شرح داده شده است .

در فصل دوم اساس طرح بهینه را شرح خواهیم داد. سپس ویژگی‌هایی از معیار D-بهینگی را بیان می‌کنیم و قضیه هم ارزی را در حالت کلی بیان کرده و در آخر فرضیه طرح های محدب برای مدل رگرسیون پواسن با ضرایب تصادفی را مورد بررسی قرار می‌دهیم .

در فصل سوم ابتدا D-بهینگی را برای مدلهایی از رگرسیون پواسن ساده حساب کرده و در آخر D-بهینگی را برای مدل رگرسیون پواسن با عرض از مبدأ تصادفی و شیب تصادفی شرح می‌دهیم.

در فصل چهارم در مورد معیار D-بهینگی برای مدل رگرسیون پواسن چندگانه با ضرایب تصادفی در دو حالت عرض از مبدا تصادفی و ضرایب تصادفی برای متغیرهای مستقل بحث کرده و برای حالت های خاصی از مدل رگرسیون پواسن دوگانه نقاط D-بهینه را حساب می‌کنیم .

فصل ۱

فصل ۱: مدل های رگرسیونی

۱-۱ مقدمه

از دیرباز نحوه ارتباط بین متغیرها مورد توجه آمارданان بوده است، مدل‌هایی که معمولاً به این داده‌ها برازش داده می‌شوند تحت عنوان مدل‌های رگرسیونی شناخته می‌شوند. در این فصل مرور کلی به انواع مدل‌های رگرسیونی خواهیم داشت. در بخش اول ساده‌ترین نوع از انواع مدل‌ها یعنی مدل‌های خطی بررسی خواهند شد. در بخش دوم مدل‌های خطی تعمیم یافته که تعمیمی از مدل‌های خطی است را شرح داده می‌شوند. در بخش سوم مدل‌های آمیخته خطی، با اضافه نمودن اثرات تصادفی به مدل خطی ارائه می‌شود در بخش چهارم تعمیم کلی از مدل‌های خطی تحت عنوان مدل‌های آمیخته خطی تعمیم یافته را شرح می دهیم و در بخش آخر ساختار مدل رگرسیون پواسن با ضرایب کلی را معرفی می‌کنیم و می‌بینیم که فرم بسته‌ای برای ماتریس اطلاع آن بدست نمی‌آید و برای حل این مشکل ماتریس شبه اطلاع را معرفی می کنیم.

۱-۲ مدل خطی

یک مدل خطی ساده‌ترین نوع ارتباط، یعنی ارتباط خطی بین متغیر پاسخ Y و مجموعه‌ای از متغیرهای توصیف‌کننده را بیان می‌کند که به شرح زیر نمایش داده می‌شود:

$$Y = \eta(\boldsymbol{x}, \boldsymbol{\beta}) = \boldsymbol{x}^T\boldsymbol{\beta} + \epsilon\ .$$

که در آن Y متغیر پاسخ، $\boldsymbol{x} = (x_1, ..., x_p)^T$ برداری از متغیرهای توصیف‌کننده، $\boldsymbol{\beta} = (\beta_0, ..., \beta_{p-1})^T$ بردار p-بعدی از پارامترهای مجهول و ϵخطای مدل با توزیع $\mathcal{N}(0, \sigma^2)$ می‌باشد.

مهم‌ترین دلیلی که استفاده از الگوهای خطی را افزایش داده است ساده بودن این مدل است به گونه‌ای‌که به راحتی می‌توان پارامترها را برآورد و تفسیر کرد در حالی‌که در الگوهای غیرخطی اغلب فرم بسته‌ای برای برآورد اثرات مدل نمی‌توان یافت.

دلیل دیگر استفاده از مدل‌های خطی این است که در مدل‌هایی که η تابعی پیچیده باشد می‌توان با استفاده از بسط تیلور،مدل را با یک الگوی خطی تقریب زد.

در حالت کلی، مدل خطی را می‌توان به فرم زیر نشان داد

$$\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{I}) \qquad (1\text{-}1)$$

یا به‌عبارت دیگر

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \ ; \ \ \boldsymbol{\epsilon} \sim \mathcal{N}(\circ, \sigma^{۲}\boldsymbol{I}) \qquad (۱-۲)$$

که $\boldsymbol{Y} = (Y_۱, ..., Y_n)^T$، $\boldsymbol{X} = (\boldsymbol{x}_۱, ..., \boldsymbol{x}_n)^T$ ماتریس $n \times p$ و معلوم است که ماتریس طرح نامیده می‌شود. به‌صورت عمومی‌تر از $\boldsymbol{f}(\boldsymbol{x}_i)$ به‌جای $\boldsymbol{x}_i$ استفاده می‌شود که در آن $\boldsymbol{f}$ تابعی از $\boldsymbol{x}$است، بنابراین خواهیم داشت

$$\boldsymbol{Y} = \boldsymbol{F}\boldsymbol{\beta} + \boldsymbol{\epsilon} \ ; \ \ \boldsymbol{\epsilon} \sim \mathcal{N}(\circ, \sigma^{۲}\boldsymbol{I})$$

که $\boldsymbol{F} = (\boldsymbol{f}(\boldsymbol{x}_۱), ...\boldsymbol{f}(\boldsymbol{x}_n))^T$.

۱-۲-۱ برآورد پارامترها

تحت فرض نرمال بودن می‌توان با استفاده از درستنمایی ماکسیمم، برآورد پارامترها را به‌دست آورد. با توجه به(۱-۱)تابع درستنمایی عبارت است از:

$$\begin{aligned} L(\boldsymbol{\beta}, \sigma^{۲}) &= f(\boldsymbol{y}; \boldsymbol{\beta}, \sigma^{۲}) \\ &= \frac{۱}{(۲\pi)^{\frac{n}{۲}} |\sigma^{۲}\boldsymbol{I}|^{\frac{n}{۲}}} e^{-\frac{۱}{۲}(\boldsymbol{y}-\boldsymbol{F}\boldsymbol{\beta})^T(\sigma^{۲}\boldsymbol{I})^{-۱}(\boldsymbol{y}-\boldsymbol{F}\boldsymbol{\beta})} \\ &= \frac{۱}{(۲\pi\sigma^{۲})^{\frac{n}{۲}}} e^{-\frac{۱}{۲\sigma^{۲}}(\boldsymbol{y}-\boldsymbol{F}\boldsymbol{\beta})^T(\boldsymbol{y}-\boldsymbol{F}\boldsymbol{\beta})} \end{aligned}$$

راه ساده‌تر این است که از لگاریتم تابع درستنمایی استفاده شود که به آن تابع امتیاز [۱] گفته می‌شود و به صورت زیر حاصل خواهد شد:

$$l = \ln L(\boldsymbol{\beta}, \sigma^{۲}) = -\frac{n}{۲} \ln ۲\pi - \frac{n}{۲} \ln \sigma^{۲} - \frac{۱}{۲\sigma^{۲}}(\boldsymbol{y} - \boldsymbol{F}\boldsymbol{\beta})^T(\boldsymbol{y} - \boldsymbol{F}\boldsymbol{\beta})$$

اگر تابع امتیاز را برابر صفر قرار دهیم آن‌گاه نقاط اکسترمم تابع برابر است با

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{F}^T\boldsymbol{F})^{-۱}\boldsymbol{F}^T\boldsymbol{Y} \quad , \quad \hat{\sigma}^{۲} = \frac{۱}{n}(\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{\beta})^T(\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{\beta}) \ .$$

که این نقاط همان مقادیر ماکسیمم تابع یا به عبارتی برآوردگرهای درستنمایی $\boldsymbol{\beta}, \sigma^{۲}$ می‌باشند.
اگر $\boldsymbol{F}^T\boldsymbol{F}$ وارون‌پذیر باشد آن‌گاه برآوردگر $\hat{\boldsymbol{\beta}}$ بهترین برآوردگر ناااریب خطی است، بهترین یعنی دارای مینیمم واریانس در میان همه برآوردگرهای ناااریبی که ترکیب خطی از مشاهدات می‌باشند. واریانس $\hat{\boldsymbol{\beta}}$ برابر است با

$$V(\hat{\boldsymbol{\beta}}) = \sigma^{۲}(\boldsymbol{F}^T\boldsymbol{F})^{-۱} \ .$$

[۱] Score function

برای این‌که وارون $\boldsymbol{F}^T\boldsymbol{F}$ موجود باشد باید $\boldsymbol{F}$ رتبه کامل باشد، یعنی $(rank(\boldsymbol{F})=p)$. اما اگر رتبه $\boldsymbol{F}$ کمتر از p باشد آن‌گاه $\boldsymbol{F}^T\boldsymbol{F}$، منفرد است در این حالت برآوردگر یکتایی برای $\boldsymbol{\beta}$ وجود ندارد و

$$\hat{\boldsymbol{\beta}}_{\circ}=(\boldsymbol{F}^T\boldsymbol{F})^{-}\boldsymbol{F}^T\boldsymbol{Y}$$

که $(\boldsymbol{F}^T\boldsymbol{F})^{-}$وارون تعمیم‌یافته است.

عدم یکتایی $\hat{\boldsymbol{\beta}}_{\circ}$ به این دلیل است که وارون تعمیم‌یافته منحصر به فرد نیست، اما می‌توان نشان داد که $\boldsymbol{F}(\boldsymbol{F}^T\boldsymbol{F})^{-}\boldsymbol{F}^T$ پایا است بنابراین

$$\boldsymbol{F}\hat{\boldsymbol{\beta}}_{\circ}=\boldsymbol{F}(\boldsymbol{F}^T\boldsymbol{F})^{-}\boldsymbol{F}^T\boldsymbol{Y}$$

یک برآوردگر یکتا برای امید متغیر پاسخ است.

۱-۳ مدل خطی تعمیم‌یافته

مدل خطی تعمیم‌یافته را می‌توان یک تعمیم انعطاف‌پذیر از رگرسیون خطی تلقی نمود که به متغیر پاسخ اجازه می‌دهد توزیع دیگری غیر از توزیع نرمال داشته باشد. نلدر[۱] و ودربرن[۲] (۱۹۷۲) مدل‌های خطی تعمیم‌یافته را به عنوان روشی برای یکسان‌سازی مدل‌های آماری مختلف شامل رگرسیون خطی، رگرسیون لجستیک و رگرسیون پواسن ارائه دادند.

۱-۳-۱ ساختار مدل

هرمدل خطی تعمیم‌یافته از سه مؤلفه تشکیل شده است: ۱- مؤلفه‌ی تصادفی، ۲- مؤلفه‌ی سیستماتیک و ۳- تابع ربط(مک‌کولاق[۳] و نلدر،۱۹۹۸).

۱- مؤلفه‌ی تصادفی: مؤلفه‌ی تصادفی توزیع احتمالی متغیر پاسخ را مشخص می‌کند. این مؤلفه شامل مشاهدات y_1، $\cdots$، y_n از متغیر تصادفی Y با توزیعی از خانواده‌ی توزیع‌های نمایی است.

خانواده توزیع‌های نمایی یک مفهوم مهم در مدل‌های آماری است. این خانواده از توزیع‌ها دارای تابع چگالی یا تابع جرم احتمالی به صورت زیر می‌باشند:

$$f(y\,|\boldsymbol{\theta})=\exp\{R(\boldsymbol{\theta})T(y)+B(\boldsymbol{\theta})+C(y)\}$$

[۱] Nelder

[۲] Wedderburn

[۳] McCullagh

که در آن $B(.)$، $T(.)$، $R(.)$ و $C(.)$ توابع حقیقی مقدار و $\boldsymbol{\theta}$ پارامتر توزیع می‌باشد. همچنین می‌توانیم ساختار دیگری با در نظر گرفتن $R(\boldsymbol{\theta}) = \frac{\gamma}{a(\phi)}$، $T(y) = y$، $B(\boldsymbol{\theta}) = -\frac{b(\gamma)}{a(\phi)}$ و $C(y) = c(y, \phi)$ برای خانواده‌ی توزیع‌های نمایی به دست آوریم. در نتیجه تابع چگالی یا تابع جرم احتمال خانواده‌ی توزیع نمایی به صورت زیر می‌باشد.

$$f(y\,|\gamma, \phi) = \exp(\frac{y\gamma - b(\gamma)}{a(\phi)} + c(y, \phi))$$

که در آن $a(.)$، $b(.)$ و $c(.)$ توابع حقیقی مقدار و ϕ پارامتر مزاحم است.

۲- مؤلفه‌ی سیستماتیک: مؤلفه‌ی سیستماتیک شامل یک بردار η است، که یک مجموعه از ترکیب‌های خطی از مجموعه متغیرهای کمکی می‌باشد؛ یعنی

$$\boldsymbol{\eta} = \boldsymbol{F\beta}$$

که در آن $\boldsymbol{F}$ ماتریس طرح نامیده می‌شود و شامل مقادیر متغیرهای کمکی و $\boldsymbol{\beta}$ یک بردار از پارامترهای مدل است. بنابراین $\boldsymbol{\eta}$ یک بردار شامل مجموعه‌ای از توابع خطی از $\boldsymbol{\beta}$ می‌باشد.

۳- تابع ربط: تابع ربط ارتباط بین مؤلفه‌ی سیستماتیک و مقادیر مورد انتظار مؤلفه‌های تصادفی را توصیف می‌کند. تابع ربط، تابعی مشتق‌پذیر و یکنوا است که میانگین پاسخ را به مؤلفه‌ی سیستماتیک پیوند می‌دهد. اگر $\mu_i = E(Y_i)$ $(i = ۱, ۲, \cdots, n)$، آن‌گاه $\eta_i = g(\mu_i)$، که در آن g تابع ربط است. در نتیجه‌می‌توانیم $g(\mu_i)$ را به صورت زیر در نظر بگیریم.

$$g(\mu_i) = \boldsymbol{f}^T(\boldsymbol{x}_i)\boldsymbol{\beta} = \sum_j x_{i,j}\beta_j \qquad\qquad j = ۱, ۲, \cdots, p$$

که در آن $\boldsymbol{f}(\boldsymbol{x}_i)$ بردار رگرسور سطر iام ماتریس طرح $\boldsymbol{F}$ و $\boldsymbol{\beta} = (\beta_۰, \beta_۲, \cdots, \beta_{p-۱})$ بردار پارامترهای مدل می‌باشند. ساده‌ترین تابع ربط، تابع ربط همانی نام دارد که در آن $g(\mu_i) = \mu_i$ به طوری که

$$\mu_i = \sum_{j=۱}^{p} x_{i,j}\beta_j.$$

۱-۳-۲ برآورد پارامترها

با توجه به خانواده توزیع‌های نمایی لگاریتم درستنمایی برابر است با:

$$l = l(\boldsymbol{\gamma}, \phi; \boldsymbol{y}) = \sum_{i=۱}^{n}(y_i\gamma_i - b(\gamma_i))/a(\phi) + \sum_{i=۱}^{n} c(y_i, \phi) \qquad (۳-۱)$$

که γ تابعی از میانگین است. برای به‌دست آوردن معادلات درستنمایی ابتدا چند معادله مفید را بیان می‌کنیم، تحت شرایط نظم داریم که:

$$E(\frac{\partial l}{\partial \gamma_i}) = ۰ \qquad (۴-۱)$$

$$var(\frac{\partial l}{\partial \gamma_i}) = -E(\frac{\partial^۲ l}{\partial \gamma_i^۲}) \qquad (۵-۱)$$

با جایگذاری لگاریتم درستنمایی(۱-۳)در روابط (۱-۴)و(۱-۵)خواهیم داشت:

- $$\frac{\partial l}{\partial \gamma_i} = (y_i - b'(\gamma_i))/a(\phi) \;\Rightarrow\; E(Y_i) = \mu_i = b'(\gamma_i)$$
- $$var(\frac{\partial l}{\partial \gamma_i}) = a^{-2}(\phi)var(Y_i)\ ,\ \ -E(\frac{\partial^2 l}{\partial \gamma_i^2}) = a^{-1}(\phi)b''(\gamma_i)$$
$$\Rightarrow var(Y_i) = a(\phi)b''(\gamma_i) = a(\phi)\upsilon(\mu_i)$$

که b' و b'' مشتق اول و دوم $b(.)$ می‌باشند. $\upsilon(\mu_i) = b''(\gamma_i)$ تابع واریانس نامیده می‌شود که نشان‌دهنده ارتباط بین میانگین و واریانس است.

همچنین با توجه به ساختار مدل‌های خطی تعمیم‌یافته خواهیم داشت

- $$\frac{\partial \gamma_i}{\partial \mu_i} = (\frac{\partial \mu_i}{\partial \gamma_i})^{-1} = (\frac{\partial^2 b(\gamma_i)}{\partial \gamma_i^2})^{-1} = \frac{1}{\upsilon(\mu_i)}$$
- $$\frac{\partial \mu_i}{\partial \boldsymbol{\beta}} = \frac{\partial \mu_i}{\partial g(\mu_i)}.\frac{\partial g(\mu_i)}{\partial \boldsymbol{\beta}} = (g_\mu(\mu_i))^{-1}\boldsymbol{f}(x_i)$$

حال می‌توان معادلات درستنمایی برای $\boldsymbol{\beta}$ رابه‌دست آورد

$$\begin{aligned}\frac{\partial l}{\partial \boldsymbol{\beta}} &= \frac{1}{a(\phi)}\sum(y_i\frac{\partial \gamma_i}{\partial \boldsymbol{\beta}} - \frac{\partial b(\gamma_i)}{\partial \boldsymbol{\beta}})\\ &= \frac{1}{a(\phi)}\sum\frac{(y_i - \mu_i)}{\upsilon(\mu_i)g_\mu(\mu_i)}\boldsymbol{f}(x_i)\\ &= \frac{1}{a(\phi)}\sum(y_i - \mu_i)w_i g_\mu(\mu_i)\boldsymbol{f}(x_i)\end{aligned}$$

که $w_i = (\upsilon(\mu_i)g_\mu^2(\mu_i))^{-1}$.

$$\frac{\partial l}{\partial \boldsymbol{\beta}} = \frac{1}{a(\phi)}\boldsymbol{F}^T\boldsymbol{W}\boldsymbol{\Delta}(\boldsymbol{y} - \boldsymbol{\mu})$$

که $\boldsymbol{W} = diag\{w_1, ..., w_n\}$ و $\boldsymbol{\Delta} = diag\{g_\mu(\mu_i)\}$.

بنابراین معادلات ماکسیمم درستنمایی برابر است با:

$$\boldsymbol{F}^T\boldsymbol{W}\boldsymbol{\Delta}\boldsymbol{y} = \boldsymbol{F}^T\boldsymbol{W}\boldsymbol{\Delta}\boldsymbol{\mu}$$

به‌گونه‌ای‌که $\boldsymbol{\Delta}, \boldsymbol{\mu}$ و $\boldsymbol{W}$ توابعی از پارامتر مجهول $\boldsymbol{\beta}$ هستند.

در بسیاری از موارد این توابع، توابعی غیر خطی از $\boldsymbol{\beta}$ هستند و نمی‌توان معادلات درستنمایی را به‌صورت آنالیزی حل کرد و باید از روش‌های عددی استفاده کرد. الگوریتم امتیازدهی فیشر یک حل عددی برای برآورد $\boldsymbol{\beta}$ ارائه می‌دهد که شکلی از روش نیوتن رافسون است(مک‌کولاک و سیرل[1]،۲۰۰۱).

[1] Searle

برای نمونه‌های بزرگ برآورد ماکسیمم درستنمایی$\boldsymbol{\beta}$،یعنی $\hat{\boldsymbol{\beta}}$، دارای واریانس به شرح زیر است که برابر معکوس ماتریس اطلاع فیشر می‌باشد :

$$var(\hat{\boldsymbol{\beta}}) = a(\phi)(\boldsymbol{F}^T\boldsymbol{W}\boldsymbol{F})^{-۱} \ .$$

۱-۴ مدل‌های آمیخته خطی

استقلال داده‌ها و فرض نرمال بودن زیر بنای اکثر روش‌های آماری متداول است. با لحاظ کردن این مفروضات بسیاری از مسائل آماری قابل تحلیل هستند. اما در بسیاری از مواقع با شرایطی روبرو می‌شویم که این مفروضات نقض می‌شوند. طیف وسیعی از داده‌ها شامل داده‌های خوشه‌ای از قبیل داده‌های طولی، داده‌های چند سطحی، اندازه‌های تکراری و ... هستند، که در آن‌ها میان مشاهدات هر واحد آزمایشی همبستگی وجود دارد که مدل‌های آمیخته خطی یک ابزار قدرتمند برای تحلیل این داده‌ها فراهم می‌آورد. در این مدل با در نظر گرفتن یک متغیر تصادفی برای هر واحد آزمایشی قادر به توضیح همبستگی بین واحدها خواهیم بود.

۱-۴-۱ مدل کلی

یک مدل آمیخته خطی را می‌توان به صورت زیر بیان کرد:

$$\boldsymbol{Y}_i = \boldsymbol{F}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i \qquad (۱-۶)$$

که $\boldsymbol{Y}_i$ بردار پاسخ $n_i \times ۱$ برای iامین موضوع $(i = ۱, ..., m)$، $\boldsymbol{F}_i$ ماتریس طرح $n_i \times p$ برای اثرات ثابت، $\boldsymbol{\beta}$بردار p-بعدی از اثرات ثابت، $\boldsymbol{Z}_i$ ماتریس طرح $n_i \times q$ برای اثرات تصادفی و $\boldsymbol{b}_i$ بردار q-بعدی از اثرات تصادفی دارای توزیع نرمال با میانگین صفر و ماتریس واریانس کواریانس $\boldsymbol{D}$ است، هم‌چنین فرض کنید برای موضوع‌های مختلف اثرهای تصادفی ناهمبسته‌اند(لرد[1] و وار[2]،۱۹۸۲). $\boldsymbol{\epsilon}_i$خطای مدل دارای توزیع نرمال با میانگین صفر و ماتریس واریانس کواریانس $\boldsymbol{R}_i$ می‌باشند که از هم مستقل‌اند، در بیشتر کاربردهای عملی $\boldsymbol{R}_i = \sigma^۲\boldsymbol{I}_i$در نظر گرفته می‌شود. با توجه به مدل، ساختار میانگین و ماتریس واریانس کواریانس $\boldsymbol{Y}_i$ به‌صورت زیر است:

$$\begin{aligned} E(\boldsymbol{Y}_i) &= \boldsymbol{F}_i\boldsymbol{\beta} \\ var(\boldsymbol{Y}_i) &= \boldsymbol{Z}_i\boldsymbol{D}\boldsymbol{Z}_i^T + \boldsymbol{R}_i \ . \end{aligned}$$

فرض شده است که $\boldsymbol{b}_i$ و $\boldsymbol{\epsilon}_i$ از هم مستقل‌اند. اگر $\boldsymbol{Y} = (\boldsymbol{Y}_۱^T, ..., \boldsymbol{Y}_m^T)^T$ بردار کل مشاهدات باشد داریم

$$\boldsymbol{V} = var(\boldsymbol{Y}) = diag(V_۱, ..., \boldsymbol{V}_m) \ .$$

[1] Laird

[2] Ware

۱-۴-۲ برآورد اثرات ثابت برای V معلوم

با توجه به اینکه $\boldsymbol{Y} \sim N(\boldsymbol{F\beta}, \boldsymbol{V})$ می‌باشد، بنابراین لگاریتم درستنمایی به‌صورت

$$l = -\frac{۱}{۲}(\boldsymbol{y} - \boldsymbol{F\beta})^T \boldsymbol{V}^{-۱}(\boldsymbol{y} - \boldsymbol{F\beta}) - \frac{۱}{۲}\log|\boldsymbol{V}| - \frac{n}{۲}\log ۲\pi$$

می‌باشد. سپس داریم

$$\frac{\partial l}{\partial \boldsymbol{\beta}} = \boldsymbol{F}^T \boldsymbol{V}^{-۱}(\boldsymbol{y} - \boldsymbol{F\beta})$$

برآوردگر ماکسیمم درستنمایی برابر است با

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{F}^T \boldsymbol{V}^{-۱}\boldsymbol{F})^{-۱}\boldsymbol{F}^T \boldsymbol{V}^{-۱}\boldsymbol{y}$$

که $\hat{\boldsymbol{\beta}}$برآوردگر یکتای $\boldsymbol{\beta}$ می‌باشد و دارای واریانس $(\boldsymbol{F}^T \boldsymbol{V}^{-۱}\boldsymbol{F})^{-۱}$ است. اگر $(\boldsymbol{F}^T \boldsymbol{V}^{-۱}\boldsymbol{F})$ وارون‌پذیر نباشد، برآورد $\boldsymbol{\beta}$ بدین شکل می‌باشد

$$\boldsymbol{\beta}_{۰} = (\boldsymbol{F}^T \boldsymbol{V}^{-۱}\boldsymbol{F})^{-}\boldsymbol{F}^T \boldsymbol{V}^{-۱}\boldsymbol{y}$$

که یکتا نیست، اما $\boldsymbol{F\beta}_{۰}$ یک برآورد یکتا برای میانگین پاسخ است زیرا $\boldsymbol{F}(\boldsymbol{F}^T \boldsymbol{V}^{-۱}\boldsymbol{F})^{-}\boldsymbol{F}^T$ به‌ازای هر $(\boldsymbol{F}^T \boldsymbol{V}^{-۱}\boldsymbol{F})^{-}$ پایا است.

۱-۴-۳ برآورد اثرات ثابت برای V نامعلوم

اگر $\boldsymbol{V} = \boldsymbol{V}(\varphi)$ باشد که در آن φ بردار مجهول ماتریس واریانس کواریانس است، آن‌گاه با توجه به قسمت قبل برآورد $\boldsymbol{\beta}$ برابر است با

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{F}^T \hat{\boldsymbol{V}}^{-۱}\boldsymbol{F})^{-۱}\boldsymbol{F}^T \hat{\boldsymbol{V}}^{-۱}\boldsymbol{y}$$

که $\hat{\boldsymbol{V}} = \boldsymbol{V}(\hat{\varphi})$ برآورد ماکسیمم درستنمایی $\boldsymbol{V}$ می‌باشد.

معادلات درستنمایی برای $\boldsymbol{V} = \boldsymbol{V}(\varphi)$ به‌صورت زیر فراهم می‌شود

$$\frac{\partial l}{\partial \varphi_k} = -\frac{۱}{۲}\left[tr(\boldsymbol{V}^{-۱}\frac{\partial \boldsymbol{V}}{\partial \varphi_k}) - (\boldsymbol{y} - \boldsymbol{F\beta})^T \boldsymbol{V}^{-۱}\frac{\partial \boldsymbol{V}}{\partial \varphi_k}\boldsymbol{V}^{-۱}(\boldsymbol{y} - \boldsymbol{F\beta})\right].$$

φ_k ، kامین عنصرφ می‌باشد.