



انتشارات، شماره ۴۴۵

الگوهای خطی تعمیر یافته

با کاربردهای آن در علوم و مهندسی

تالیف:

ریموند اچ. میرز

دکلاس سی. منگمری

جی. جنوفری وینینگ

ترجمه:

دکتر حسینعلی نیرومند

مایرز، ریموند، ۱۹۲۷ - م.
الگوهای خطی تممیم یافته با کاربردهای آن در علوم و مهندسی / تألیف ریموند اچ، میرز،
داگلاس سی. منتگمری، جی. جنو فری وینینگ، ترجمه حسینعلی نیرومند. مشهد: دانشگاه فردوسی
مشهد، ۱۳۸۴.

۴۱۶ ص. - (انتشارات دانشگاه فردوسی مشهد؛ ۴۴۵)

ISBN: 964-386-099-x

۲۴۰۰۰ ریال

فهرست نویسی بر اساس اطلاعات فیبا.
عنوان اصلی: Generalized linear models with applications in engineering and the sciences.

کتابنامه: ص. ۴۰۱-۴۰۵.

۱. الگوهای خطی (آمار). الف. منتگمری، داگلاس، Montgomery, Douglas c.
ب. وینینگ، جفری، ۱۹۵۴ - م. Vining G. Geoffery، ج. نیرومند، حسینعلی،
۱۳۲۶ - مترجم. د. دانشگاه فردوسی مشهد، ه. عنوان.

۵۱۹/۵۳۵

۷ الف ۲ م / QA ۲۷۶

۱۳۸۴

۸۴-۲۳۷۶۳ م

کتابخانه ملی ایران



انتشارات، شماره ۴۴۵

الگوهای خطی تممیم یافته با کاربردهای آن در مهندسی و علوم

تألیف

ریموند اچ. میرز، داگلاس سی. منتگمری، جی. جنو فری وینینگ

ترجمه

دکتر حسینعلی نیرومند

ویراستار علمی

دکتر حسنعلی آذرنوش

وزیری، ۴۱۶ صفحه، ۱۰۰۰ نسخه، چاپ اول، پاییز ۱۳۸۴

امور فنی و چاپ: مؤسسه چاپ و انتشارات دانشگاه فردوسی مشهد

بها: ۲۴۰۰۰ ریال

فهرست

۷	پیشگفتار مؤلفان
۱۱	فصل اول مقدمه‌های بر الگوهای خطی تعمیم یافته.....
۱۱	۱-۱ الگوهای خطی.....
۱۳	۲-۱ الگوهای غیر خطی.....
۱۵	۳-۱ الگوی خطی تعمیم یافته.....
۱۹	فصل دوم الگوهای رگرسیون خطی.....
۱۹	۱-۲ الگوی رگرسیون خطی و کاربرد آن.....
۲۰	۲-۲ الگوهای رگرسیون چندگانه.....
۴۷	۳-۲ برآورد پارامتر با استفاده از درست نمایی ماکزیمم.....
۵۰	۴-۲ بررسی درستی الگو.....
۶۷	۵-۲ برآورد پارامتر با کمترین مربعات وزنی.....
۷۲	تمرینها.....
۸۵	فصل سوم الگوهای رگرسیون غیرخطی.....
۸۵	۱-۳ الگوهای رگرسیون خطی و غیرخطی.....
۸۸	۲-۳ تبدیل کردن به یک الگوی خطی.....
۹۲	۳-۳ برآورد پارامتر در یک سیستم غیرخطی.....
۱۱۰	۴-۳ استنباط آماری در رگرسیون غیرخطی.....
۱۱۵	۵-۳ رگرسیون غیرخطی وزنی.....

- ۳-۶ مثالهایی از الگوهای رگرسیون غیرخطی..... ۱۱۶
 تمرینها..... ۱۱۷

فصل چهارم الگوهای رگرسیون لجستیک و پواسن..... ۱۲۵

- ۴-۱ الگوهای رگرسیونی که واریانس تابعی از میانگین است..... ۱۲۵
 ۴-۲ الگوی رگرسیون لجستیک..... ۱۲۶
 ۴-۳ برآورد پارامتر با استفاده از درست نمایی ماکزیمم..... ۱۲۸
 ۴-۴ اشکال مختلف استنباط آماری با استفاده از رگرسیون لجستیک..... ۱۳۲
 ۴-۵ مثالهایی از کاربرد رگرسیون لجستیک..... ۱۴۱
 ۴-۶ بررسیهای دیگر در رگرسیون لجستیک..... ۱۴۸
 ۴-۷ مفهوم پراکندگی اضافی در رگرسیون لجستیک..... ۱۵۵
 ۴-۸ مقدمه‌های بر رگرسیون پواسن..... ۱۶۱
 ۴-۹ برآوردگر درست‌نمایی ماکزیمم برای رگرسیون پواسن..... ۱۶۲
 ۴-۱۰ کاربردهای رگرسیون پواسن..... ۱۶۴
 ۴-۱۱ چند مثال از کاربرد رگرسیون پواسن..... ۱۶۶
 ۴-۱۲ متغیرهای رده بندی و تعمیم آن به الگوی ANOVA..... ۱۸۲
 تمرینها:..... ۱۸۳

فصل پنجم خانواده الگوهای خطی تعمیم یافته..... ۱۹۱

- ۵-۱ خانواده نمایی توزیعها..... ۱۹۱
 ۵-۲ ساختار رسمی رده الگوهای خطی تعمیم یافته..... ۱۹۵
 ۵-۳ معادلات درست‌نمایی برای الگوهای خطی تعمیم یافته..... ۱۹۷
 ۵-۴ شبه درست‌نمایی..... ۲۰۱
 ۵-۵ توزیعهای مهم دیگر برای الگوهای خطی تعمیم یافته..... ۲۰۳
 ۵-۶ ردهایی از توابع پیوند - توابع توان..... ۲۰۷
 ۵-۷ استنباط و تحلیل مانده برای الگوهای خطی تعمیم یافته..... ۲۰۹
 ۵-۸ چند مثال از توزیع گاما..... ۲۱۲
 تمرینها:..... ۲۲۷

فصل ششم	معادلات برآورد تعمیم یافته	۲۳۳
۱-۶	طرح ریزی داده ها برای مطالعات طولی	۲۳۴
۲-۶	استفاده از ماتریس همبستگی R	۲۳۵
۳-۶	شیوه تکراری در حالت نرمال و پیوند همانی	۲۳۷
۴-۶	معادلات برآورد تعمیم یافته برای الگوهای خطی تعمیم یافته تر	۲۳۹
۵-۶	مثالها	۲۴۷
۶-۶	خلاصه	۲۷۴
	تمرینها	۲۷۴

فصل هفتم	دیگر پیشرفته‌ها و کاربردهای GLM	۲۸۳
۱-۷	مقدمه	۲۸۳
۲-۷	طرحهای آزمایشی برای الگوهای خطی تعمیم یافته	۲۸۴
۳-۷	کیفیت نتایج مجانبی و موارد وابسته به آن	۳۰۵
۴-۷	تحلیل GLM آزمایشهای غربال شده	۳۲۴
۵-۷	GLM و تبدیل داده ها	۳۳۹
۶-۷	الگوسازی میانگین و واریانس فرآیند با استفاده از GLM	۳۴۹
۷-۷	الگوهای جمعی تعمیم یافته	۳۶۱
	تمرینها	۳۷۴

نمایه		۳۹۹
مراجع		۴۰۱

پیشگفتار مؤلفان

کتاب حاضر؛ کتاب درسی مقدماتی در الگوی خطی تعمیم یافته یا *GLM* است. این کتاب برای کسانی که در تحلیل رگرسیون دوره‌ای را گذرانده و با برازش مقدماتی الگو و شیوه‌های استنباط آماری که معمولاً در سالهای اول کارشناسی ارشد یاد می‌گیرند، آشنا باشند مناسب است. داشتن زمینه قبلی در برآورد درست نمایی ماکزیمم برای الگوهای رگرسیون خطی و در رگرسیون غیر خطی مفید خواهد بود، و این موضوعات با جزئیات بیشتری در فصلهای ۲ و ۳ مورد بحث قرار خواهد گرفت.

کتاب چند جنبه منحصر به فرد دارد: ابتدا طرز عمل کاملی از رگرسیون لجستیک و پواسن را ارائه می‌کند. درحالی‌که حالات خاصی از *GLM* هستند، حالت‌های بسیار مهمی هستند و در خور توجه بیشتری می‌باشند. نحوه عملکرد این دو حالت به خوبی خواننده را با *GLM* آشنا می‌کند. دوم این که، مقدمه‌ای از معادلات برآورد تعمیم یافته را فراهم می‌کند که این موضوع ارتباط نزدیکی با *GLM* دارد و از اهمیت فزاینده‌ای برخوردار است. ضمن این که بسیاری از کاربردهای معادلات برآورد تعمیم یافته به زمینه‌های زیستی، در ارتباط با داده‌های طولی است، ولی این موضوع جذابیت رو به رشدی در علوم فیزیکی و مهندسی دارد که کاربردهای آن در مورد مشاهدات همیشه بیش می‌آید.

جنبه مفید دیگر این کتاب مثال‌های زیادی در *GLM* است که دامنه آن از کاربردهای «کلاسیک» در علوم زیستی و علوم مربوط به دارویی زیستی تا مهندسی، تضمین کیفی و آزمایش‌های طرح شده است. ما معتقدیم که دامنه موضوعات و کاربردها هم برای تمرین افراد حرفه‌ای در زمینه‌های گوناگون و هم به عنوان یک کتاب درسی برای دومین دوره در رگرسیون به این کتاب جاذبه زیادی می‌دهد. سرانجام راهنمایی و هدایت قابل ملاحظه‌ای برای محاسبه کردن فراهم می‌شود. در سرتاسر کتاب از *SAS PROC GENMOD* و *S-PLUS* برای تشریح و بیان هر جنبه برازش الگو، استنباط و بررسی درستی تشخیص استفاده می‌شود. در دسترس بودن وسیع این بسته‌های نرم افزار این کتاب را عملاً مفید می‌سازد.

سعی ما بر این است که این کتاب به عنوان یک کتاب درسی مقدماتی برای طیف وسیعی از مخاطبان قابل دسترس باشد. کارهای عالی دیگری در *GLM* وجود دارد ولی بسیاری از این

کارها در سطح خیلی بالاتر بوده و ابتدائاً برای آماردانان نظری نوشته شده یا این کارها در مورد یک موضوع بوده و مفاد آن خیلی کمتر از کتاب حاضر است. هدف کتاب، بسترسازی برای گروه وسیعی از مهندسان، دانشمندان و آماردانان است تا نظریه *GLM* و چگونگی کاربری آن را یاد بگیرند.

در فصل اول مقدمه‌ای آمده تا به مخاطب به عنوان یک خواننده دیدگاهی بر این مبنی که چه نوع مسائلی به استفاده از الگوی خطی تعمیم یافته کمک می‌کند، ارائه دهد. در فصل دوم بسیاری از مفاهیم مقدماتی رگرسیون خطی چندگانه مرور می‌شود. مفاهیم بنیادی نظیر روش برآورد کمترین مربعات و درست‌نمایی ماکزیمم مورد بحث قرار می‌گیرد. همچنین برآورد فواصل اطمینان، روشهای آزمون فرضیه و تکنیکهای بررسی درستی تشخیص الگو مانند نمودار مانده‌ها و تشخیصهای مؤثر ارائه می‌شود. احتمالاً این مطالب برای خواننده یا دانشجویان علاقه‌مند به الگوی خطی تعمیم یافته بیشتر یک مرور خواهد بود. این فصل بحث روشهایی برای به کارگیری وضعیتهایی که در آن فرض همگنی واریانس نقض می‌گردد و منجر به کمترین مربعات وزنی و تبدیل داده‌ها می‌شود به پایان می‌رسد. برای خواننده بخصوص مهم است که این ایده‌ها را فرا بگیرد، زیرا الگوهای خطی تعمیم یافته در وضعیتهایی مورد استفاده قرار می‌گیرد که به خاطر ماهیت توزیع پاسخ، واریانس ثابت نیست.

فصل سوم ارائه مختصری از ایده‌های بنیادی رگرسیون غیر خطی است. در این فصل مطلب را با طرح تمایز آشکاری که بین الگوهای خطی و غیر خطی وجود دارد آغاز می‌کنیم. این موضوع با بررسی کاملاً معقولانه‌ای از برآورد پارامتر در الگوهای غیر خطی همراه است. برآورد کمترین مربعات و برآورد درست‌نمایی ماکزیمم هر دو بحث می‌شوند. مواد این فصل هنگامی که *GLM* را مطالعه می‌کنیم بی‌نهایت با اهمیت است، زیرا ماهیت اکثر الگوهای خطی تعمیم یافته غیر خطی است. مثالهایی برای تشریح برازش و استنباط مربوط به الگوهای غیر خطی ارائه می‌شود.

در فصل چهارم دو حالت خاص و مهم الگوی خطی تعمیم یافته یعنی رگرسیون لجستیک و پواسن را توسعه داده و تشریح می‌کنیم. الگوهای رگرسیون پواسن و لجستیک به‌طور گسترده در علوم زیستی، پزشکی و اجتماعی مورد استفاده قرار می‌گیرد و استفاده فزاینده‌ای در علوم فیزیکی و مهندسی دارد. توسعه نظری در این فصل پایه‌ای را به وجود می‌آورد که با آن فهم *GLM* خیلی ساده‌تر می‌شود. مثالهای متنوعی در زمینه‌های مختلف ارائه می‌شود که هم مطالعات مشاهده‌شدنی و هم آزمایشهای طرح شده را تشریح می‌کند.

در فصل پنجم خانواده نسایی توزیعها معرفی می‌شود و فرآیند الگوسازی را برای یک الگوی خطی تعمیم یافته با استفاده از یک تابع پیوند مهیج می‌کند. رابطه بین کمترین مربعات وزنی و درست نمایی ماکزیمم را توسعه داده و از آن به عنوان یک وسیله محاسباتی برای برآورد پارامتر در *GLM* استفاده می‌کنیم. «خانواده گاما» را بررسی نموده و با به کارگیری *SAS*، *PROC GENMOD* و *S-PLUS*، مطالب مختلفی را ارائه می‌کنیم.

فصل ششم به معادلات برآورد تعمیم یافته (*GEE*ها) اختصاص یافته است. این معادلات وقتی ساختارهای همبستگی داخل واحدهای آزمایشی بزرگ که در آن مشاهدات «تکرار شده» داریم تعمیمهای *GLM* هستند. استفاده *GEE*ها با انواع خاصی از طرحهای آزمایشی صنعتی و کاربردهای طب زیستی را بحث کرده که آنها را با مثال تشریح می‌کنیم.

سرانجام فصل هفتم مجموعه‌ای از موضوعات در مورد *GLM*ها و *GEE*ها را شامل می‌شود و دیدگاه بهتری در زمینه آزمایشهای طرح شده و *GLM* را ارائه می‌دهد؛ از جمله استفاده از نمودارهای مانده خاص و نمودارهای اثری که تعمیمهای طبیعی آنها هستند که با تحلیل نظریه نرمال قراردادی آزمایشهای عاملی بکار می‌روند. کارایی طرحهای استاندارد در حالت *GLM*ها و *GEE*ها بحث و تشریح می‌گردد. همچنین کارآ بودن *GLM* به عنوان جانشینی برای تبدیلات داده‌ها را بحث می‌کنیم و مقدمه مختصری از الگوهای جمعی تعمیم یافته نیز ارائه می‌گردد.

در طول بهار سال ۲۰۰۱ مؤلفان از پیش نویس (طرح اجمالی) این کتاب در تدریس دوره‌های سطح پیشرفته الگوهای خطی تعمیم یافته در ویرجینیا *TeCh* و آریزونا *State*، استفاده کرده اند، که برای دانشجویان در این دوره‌ها بازخورد و انتقاد با ارزشی فراهم گردید و کار و کوشش سخت آنها نقش عمده‌ای در شکل گیری جنبه نهایی این کتاب ایفاء نمود. از کتینا اسکینر، دونگپورن جیرکبا پورن و دکتر شارون لوینر به خاطر کمک ارزشمندشان در محاسبات متشکر و سپاسگزاریم.

ریموند. اچ. میرز
داکلاس سی مونتگمری
جی. جوزفری وینینگ