



تحلیل بیزی مدل رگرسیون چندکی تعمیم یافته

افشین فلاح*

دانشگاه بین‌المللی امام خمینی (ره)

منیر میرزایی

دانشگاه بین‌المللی امام خمینی (ره)

چکیده

در این مقاله برای تحلیل رگرسیون چندکی تعمیم یافته یک مدل بیزی پیشنهاد شده است. در این نوع تحلیل رگرسیونی به جای میانگین، مجموعه‌ای از چندک‌ها برای تبیین توزیع متغیر پاسخ به‌کار گرفته می‌شوند. کارایی مدل پیشنهادی در قالب یک مطالعه شبیه‌سازی مورد ارزیابی قرار گرفته است.

واژه‌های کلیدی: چندک، رگرسیون چندکی، بیز، براوردگرهایی از نوع ماکسیمم درست‌نمایی

Mathematics Subject Classification [2010]: 62F15, 62G08

۱ مقدمه

در روش‌های معمول تحلیل رگرسیونی میانگین متغیر پاسخ یا تابعی از آن به‌صورت شرطی بر مجموعه‌ای از متغیرهای تبیینی مدل‌بندی می‌شود. چون میانگین یکی از معیارهای تمرکز است، به تنهایی نمی‌تواند توزیع متغیر پاسخ را در فرایند مدل‌سازی تبیین نماید. چندک‌ها معیارهایی هستند که در کنار هم می‌توانند شکل توزیع را به‌صورت جامع‌تری به تصویر بکشند. از طرفی، هنگامی که داده‌های دورافتاده‌ای وجود داشته باشد، علاوه بر میانگین شرطی متغیر پاسخ رفتار مقادیر کرانگین این متغیر نیز از اهمیت برخوردار است. یک راه برای کاهش اثر این‌گونه داده‌ها استفاده از معیار قدرمطلق خطا است [۴]. [۲] و [۵] تحلیل رگرسیون میانه را که در آن براورد پارامترهای مدل رگرسیونی با هدف مینیمم ساختن قدرمطلق خطا صورت می‌پذیرد، برای مدل‌سازی داده‌هایی که دارای نقاط دور افتاده هستند، پیشنهاد نمودند. رگرسیون چندکی را می‌توان تعمیمی از رگرسیون میانه دانست که در آن به جای مدل‌بندی میانگین متغیر پاسخ، چندک‌های این متغیر به‌صورت شرطی بر مجموعه‌ای از متغیرهای تبیینی مدل‌بندی می‌شوند. چون الگوریتم‌های مورد استفاده برای برازش مدل‌های رگرسیون چندکی لزوماً همگرایی و جواب یکتا را تضمین نمی‌کنند، [۱] ایده‌ی رگرسیون چندکی تعمیم یافته را مطرح ساختند که مبتنی بر براوردگرهایی از نوع ماکسیمم درست‌نمایی هستند و بر اساس یک تابع تاثیر تعریف می‌شوند. در این مقاله، رویکرد بیزی تحلیل رگرسیون چندکی تعمیم یافته مورد توجه قرار گرفته است. برای این منظور، با در نظر گرفتن توزیع‌های پیشین مناسب برای پارامترها، توزیع پسین پارامترهای مدل محاسبه شده است. نشان داده شده است که به‌دلیل شکل پیچیده تابع درست‌نمایی مبتنی بر مدل رگرسیون چندکی تعمیم یافته، توزیع پسین فاقد فرم بسته بوده و شکل پیچیده‌ای دارد. از این‌رو توزیع‌های پسین شرطی کامل پارامترها محاسبه و الگوریتم گیبز مناسب برای نمونه‌گیری از توزیع پسین توسعه داده شده است. هم‌چنین کارایی مدل بیز پیشنهادی در براورد پارامترهای مدل در مقایسه با مدل فراوانی‌گرایانه بر اساس یک مطالعه شبیه‌سازی، مورد ارزیابی قرار گرفته است. در بخش ۲ تحلیل رگرسیون چندکی تعمیم یافته از دیدگاه بیزی مورد بحث قرار گرفته و نحوه تعیین توزیع پیشین، محاسبه توزیع پسین، نمونه‌گیری از توزیع پسین و غیره، شرح داده شده است. در بخش ۳ برای ارزیابی و مقایسه مدل‌های رگرسیون چندکی و رگرسیون چندکی تعمیم یافته در رهیافت فراوانی‌گرایانه و بیزی مطالعه‌ای شبیه‌سازی صورت پذیرفته است.

* سخنران