دانشگاه علامه طباطبائی
دانشکده‌ی اقتصاد
گروه آمار، ریاضی و کامپیوتر

پایان‌نامه برای دریافت درجه‌ی کارشناسی ارشد **آمار ریاضی**

عنوان

مدل‌سازی داده‌های دومتغیره‌ی همبسته و کاربرد آن در داده‌های واقعی

پژوهش‌گر

نرگس راموز

استاد راهنما

دکتر فرزاد اسکندری

استاد مشاور

دکتر محمد بامنی‌مقدم

بهمن۱۳۸۹

بسم الله الرحمن الرحيم

کلیه‌ی حقوق مادی و معنوی اعم از چاپ و تکثیر، نسخه‌برداری، ترجمه، اقتباس و ... از این پایان‌نامه

برای دانشگاه علامه طباطبایی محفوظ است. نقل مطالب با ذکر منبع مانعی ندارد.

تأیید پایان‌نامه‌ی کارشناسی ارشدتوسط دانشجو

عنوان پایان‌نامه: مدل‌سازی داده‌های دومتغیره‌ی همبسته و کاربرد آن در داده‌های واقعی

نام دانشجو: نرگس راموز

شماره‌ی دانشجویی: ۸۷۱۱۲۸۲۰۳

استاد راهنما: دکتر فرزاد اسکندری

این‌جانب **نرگس راموز** دانشجوی کارشناسی ارشد رشته‌ی **آمار ریاضی** دانشکده‌ی اقتصاد دانشگاه علامه طباطبایی گواهی می‌نمایم پژوهش‌های ارائه شده در پایان‌نامه با عنوان مذکور توسط شخص این‌جانب انجام شده است و درستی مطالب نگارش یافته مورد تأیید می‌باشد. همچنین گواهی می‌نمایم مطالب مندرج در پایان‌نامه تاکنون برای دریافت هیچ نوع مدرک یا امتیازی توسط این‌جانب یا فرد دیگری در هیچ کجا ارائه نشده است و در نگارش متن پایان‌نامه شیوه‌ی نگارش مصوب دانشکده‌ی اقتصاد را به‌طور کامل رعایت نموده‌ام. چنان‌چه در هر زمان خلاف آن‌چه گواهی نموده‌ام مشاهده گردد خود را از آثار حقیقی و حقوقی ناشی از دریافت مدرک کارشناسی ارشد محروم می‌دانم و هیچ‌گونه ادعایی نخواهم داشت.

تاریخ: **امضا دانشجو:**

تقدیم به همه ی آن هایی که

حقشان بر گردنم است

سپاس‌گزاری

سپاس خدای را که هر توفیقی در گرو عنایت اوست. اکنون که با یاری او توانسته‌ام تلاشی هر چند ناچیز را در راه کسب دانش به انجام رسانم، بر خود لازم می‌دانم از استاد راهنمای بزرگوارم، جناب آقای دکتر اسکندری، که به پایان رساندن این تحقیق جز با راهنمایی‌های دلسوزانه و هدایت‌های بی‌دریغ ایشان میسر نبود، قدردانی نمایم.

از استاد مشاورم جناب آقای دکتر بامنی مقدم که تذکراتشان باعث غنای پایان‌نامه شد، تشکر می‌نمایم. همچنین از جناب آقای دکتر نواب‌پور و جناب آقای دکتر نعمت‌الهی که زحمت داوری این اثر را به عهده داشتند سپاس‌گزارم.

از خانواده‌ام، به‌ویژه پدر و مادرم که با حمایت‌های خویش، همواره مرا پشتیبانی کرده‌اند نهایت سپاس و قدرشناسی را دارم. درپایان، از همراهی و راهنمایی‌های ریاست محترم اداره‌ی اقتصادی تعاون استان تهران و کارشناسان آن اداره که تجربیات خویش را در اختیارم قرار دادند، کمال تشکر را دارم.

امیدوارم بتوانم از عهده ادای حق این عزیزان برآیم.

بهمن۸۹

فهرست مطالب

فهرست مطالب ب

فهرست جدول‌ها ج

فهرست شکل‌ها چ

نمادها و علایم اختصاری ح

۱ کلیات پژوهش د

۱.۱ مقدمه ۱

۲.۱ مدل رگرسیون پواسون ۱

۳.۱ هدف‌های پژوهش ۴

۴.۱ فرض‌ها و پرسش‌ها ۵

۲ مبانی نظری و مروری بر مطالعه‌های انجام گرفته ۶

۱.۲ مدل‌سازی داده‌های شمارشی ۷

۱.۱.۲ مقدمه ۷

۲.۱.۲ تحلیل رگرسیونی داده‌های شمارشی ۷

۳.۱.۲ مدل رگرسیون پواسونی تک متغیره ۸

۲.۲ بیش پراکندگی در مدل رگرسیون پواسون ۱۰

۱.۲.۲ روش متغیر پنهان ۱۰

۲.۲.۲ مدل دو جمله‌ای منفی تک‌متغیره ۱۱

۳.۲.۲ مدل پواسون - گامای تک متغیره (PGA) ۱۳

۴.۲.۲ مدل پواسون - لگ نرمال (PLN) ۱۵

۱.۴.۲.۲ مدل پواسون - لگ نرمال (حالت ضربی اثرهای تصادفی) . . . ۱۵

۲.۴.۲.۲ مدل آمیخته‌ی پواسون - لگ‌نرمال با اثرهای تصادفی جمعی . . ۱۷
۳.۲ مروری برمطالعه‌های انجام شده ۱۹
۱.۳.۲ توزیع‌های شمارشی چند متغیره ۱۹
۲.۳.۲ مدل‌های داده‌های شمارشی همبسته ۲۰
۳.۳.۲ نگاهی بر روش‌های مختلف براورد مدل ۲۳
۴.۲ معرفی مدل ۲۵
۱.۴.۲ مدل پواسون - لگ نرمال دومتغیره ۲۵
۲.۴.۲ براورد پارامترهای مدل ۲۶
۵.۲ مقدمه‌ای بر تحلیل بیزی ۲۷
۱.۵.۲ توزیع پیشین و انواع آن ۲۸
۱.۱.۵.۲ انواع توزیع پیشین ۲۸
۲.۵.۲ مخاطره و تصمیم بیزی ۲۸
۳.۵.۲ روش تصمیم‌سازی بیزی ۲۹
۴.۵.۲ روش سلسله مراتبی با رهیافت بیزی ۳۰
۵.۵.۲ توزیع‌های شرطی کامل ۳۱
۶.۵.۲ زنجیر مارکوف ۳۱
۷.۵.۲ انتگرال مونت - کارلویی ۳۳
۸.۵.۲ مونت کارلوی زنجیر مارکوفی(MCMC) ۳۳
۱.۸.۵.۲ نمونه‌گیری گیبزی ۳۴
۲.۸.۵.۲ آلگوریتم متروپولیس - هستینگز ۳۴
۹.۵.۲ دوره‌ی سوخت ۳۵
۱۰.۵.۲ انتخاب توزیع پیشنهادی در آلگوریتم متروپلیس - هستینگز ۳۶
۱۱.۵.۲ مبحث‌های همگرایی ۳۸
۱.۱۱.۵.۲ معرفی همگرایی ۳۸
۲.۱۱.۵.۲ انتخاب نقطه‌ی اولیه و تعیین طول دوره‌ی مقدماتی ۳۹
۳.۱۱.۵.۲ روش‌های تشخیص همگرایی ۴۰
۱۲.۵.۲ مقایسه بیزی مدل‌ها ۴۱
۱.۱۲.۵.۲ ملاک اطلاع بیزی ۴۲
۲.۱۲.۵.۲ ملاک عامل بیزی ۴۳

۳.۱۲.۵.۲ ملاک اطلاع انحرافی ۴۴

۳ تعیین براورد پارامترهای مدل با رهیافت بیزی ۴۷

۱.۳ مقدمه ۴۸

۲.۳ مدل پواسون-لگ‌نرمال دومتغیره ۴۸

۳.۳ توزیع پسین ۵۰

۴.۳ نمونه‌گیری از توزیع پسین ۵۲

۱.۴.۳ توزیع شرطی کامل $\overrightarrow{b_i}$ ۵۲

۲.۴.۳ توزیع شرطی کامل Σ^{-1} ۵۴

۳.۴.۳ توزیع شرطی کامل β ۵۵

۴.۴.۳ توزیع شرطی کامل β_β ۵۷

۵.۴.۳ توزیع شرطی کامل V_β^{-1} ۵۸

۶.۴.۳ توزیع شرطی کامل V_Σ^{-1} ۵۹

۴ کاربرد مدل با داده‌های واقعی ۶۰

۱.۴ کاربردی از داده‌های شمارشی همبسته با داده‌های واقعی ۶۱

۱.۱.۴ مقدمه ۶۱

۲.۴ اشتغال، کارآفرینی و بنگاه‌های اقتصادی زودبازده ۶۱

۱.۲.۴ شرکت تعاونی ۶۲

۲.۲.۴ مقررات مالی شرکت تعاونی و پرداخت حق آموزش ۶۳

۳.۲.۴ مواد و روش‌ها ۶۳

۳.۴ کاربرد مدل دومتغیره‌ی همبسته شمارشی با داده‌های واقعی ۶۵

۱.۳.۴ شاخص‌های توصیفی داده‌ها ۶۵

۲.۳.۴ تشخیص مدل ۶۶

۳.۳.۴ تحلیل داده‌ها بر مبنای رهیافت بیزی ۶۶

۴.۴ براورد پارامترها ۶۸

۱.۴.۴ مدل پواسون_لگ نرمال دومتغیره ۶۸

۲.۴.۴ مدل پواسون-گامای دو متغیره ۷۲

۳.۴.۴ استنباط شهودی در مورد پارامترهای مدل ۷۳

۵.۴ مقایسه‌ی مدل‌ها و تعیین مدل مناسب ۷۵

۵ نتیجه‌گیری و پیش‌نهادها **۷۶**

۱.۵ مقدمه ۷۷

۲.۵ تفسیر نتیجه‌های حاصل از داده‌های واقعی ۷۷

۱.۲.۵ تفسیر پارامترهای مدل ۷۷

۲.۲.۵ تفسیر همبستگی موجود درمدل ۷۹

۳.۵ پیش‌نهادها ۸۲

واژه‌نامه‌ی فارسی به انگلیسی **۸۵**

پیوست الف آلگوریتم‌ها و خروجی‌های پژوهش **۹۱**

کتاب‌نامه **۹۷**

فهرست جدول‌ها

۱.۴ جدول آمارهای توصیفی ۶۵
۲.۴ جدول آزمون همبستگی بین دو متغیر پاسخ ۶۵
۳.۴ جدول آزمون ناپارامتری کولموگروف-اسمیرونوف ۶۶
۴.۴ آماره‌های خلاصه‌ی پسین پارامترهای مدل اشباع شده‌ی پواسون-لگ‌نرمال دومتغیره ۷۰
۵.۴ آماره‌های خلاصه‌ی پسین پارامترهای مدل اشباع شده‌ی پواسون-لگ‌نرمال دومتغیره ۷۱
۶.۴ آماره‌های خلاصه‌های پسین پارامترهای مدل کاهش‌یافته‌ی پواسون-لگ‌نرمال دومتغیره ۷۱
۷.۴ آماره‌های خلاصه‌ی پسین پارامترهای مدل کاهش‌یافته‌ی پواسون-لگ‌نرمال دومتغیره ۷۱
۸.۴ آماره‌های خلاصه‌ی پسین پارامترهای مدل اشباع شده‌ی پواسون-گامای دومتغیره ۷۳
۹.۴ آماره‌های خلاصه‌ی پسین پارامترهای مدل اشباع شده‌ی پواسون-گامای دومتغیره ۷۳
۱۰.۴ آماره‌های خلاصه‌ی پسین پارامترهای مدل کاهش‌یافته‌ی پواسون-گامای دومتغیره ۷۴
۱۱.۴ خلاصه‌های آماری پسین پارامترهای مدل کاهش‌یافته‌ی پواسون-لگ‌نرمال دومتغیره ۷۴
۱۲.۴ جدول ملاک مقایسه‌ی بیزی DIC ۷۵

۱.۵ براورد پارامترهای مربوط به ماتریس سیگما Σ ۸۰
۲.۵ براورد پارامترهای رگرسیون پواسون تک‌متغیره ۸۱
۳.۵ براورد پارامترهای رگرسیون پواسون تک‌متغیره ۸۱

فهرست شکل‌ها

۱.۴ مدل گرافیکی رسم شده در منوی Doodle از نرم افزار WinBUGS ۶۸

۲.۴ خروجی در WinBUGS برای مدل اشباع شده‌ی پواسون-لگ‌نرمال دو متغیره ۷۰

۳.۴ خروجی WinBUGS مربوط به نمودار *Auto Correlation* در مدل پواسون-لگ‌نرمال دو متغیره . ۷۴

۱.۵ خروجی *WinBUGS* مربوط به نمودار kernel دردensity مدل پواسون-لگ‌نرمال دومتغیره . . . ۹۳

۲.۵ خروجی *WinBUGS* در مدل کاهش یافته‌ی پواسون- گامای دومتغیره ۹۴

۳.۵ خروجی *WinBUGS* در مدل اشباع شده‌ی پواسون- گامای دو متغیره ۹۴

۴.۵ خروجی *WinBUGS* در مدل کاهش یافته‌ی پواسون-لگ‌نرمال دومتغیره ۹۵

۵.۵ خروجی SPSS مربوط به همبستگی دو متغیر پاسخ ۹۵

۶.۵ خروجی SPSS مربوط به آزمون کولموگروف-اسمیرونوف ۹۶

نمادها و علایم اختصاری

نماد	تعریف
AIC	*Akaike Information Criterion*
BIC	*Bayesian Information Criterion*
DIC	*Deviance Information Criterion*
EM	*Expectation Maximization*
GP	*Generalize Poisson distribution*
HPD	*Highest Posterior Density*
IWLS	*Iterative Weighted Least Square Proposal*
MCMC	*Markov Chain Mont Carlo method*
ML	*Maximom Likelihood*
MVPLN	*MultiVariate Poisson Log Normal*
NB	*Negative Binomial*
PGA	*Poisson Gamma*
PIG	*Poisson Inverse Gaussian*
PLN	*Poisson Log Normal*
HPD	*Highest Posterior Density*

چکیده

یکی از موضوعات کلیدی و مطرح در حوزه‌ی آمار کاربردی، مبحث داده‌های شمارشی همبسته و تشخیص مدل‌های مناسب برای تحلیل این‌گونه داده‌هاست. روش‌های محدودی برای تحلیل داده‌های شمارشی وجود دارد. استفاده از مدل رگرسیون پواسون یکی از رایج‌ترین مدل‌ها برای تحلیل این‌گونه داده‌هاست، ولی بنابر محدودیت موجود در استفاده از این مدل (شرط برابری میانگین و واریانس) و در نتیجه برای کنترل بیش‌پراکندگی در مدل مذکور، راه‌کارهایی ارائه شده است که یکی از بهترین آن‌ها استفاده از مدل‌های آمیخته می‌باشد، از این رو این مطالعه به بررسی داده‌های شمارشی دو متغیره‌ی همبسته با استفاده از مدل‌های پواسون-لگ‌نرمال و پواسون - گاما می‌پردازد. با انتخاب این مدل‌های آمیخته می‌توان محدودیت مثبت بودن همبستگی بین متغیرهای پاسخ در مدل را برطرف کرد. در این مطالعه، روش سلسله مراتبی با رهیافت بیزی برای براورد پارامترهای مدل پیشنهاد شده است. همچنین، می‌توان بازه‌های چگال‌ترین پسین را برای تک تک ضریب‌های مدل محاسبه کرد. در ادامه‌ی این پژوهش، با استفاده ازشاخص بیزی انتخاب مدل نشان داده شد که مدل پواسون-لگ نرمال دومتغیره مدل مناسب‌تری نسبت به مدل پواسون -گامای دو متغیره است. در پایان مدل ارائه شده در داده‌های واقعی مربوط به بخش اقتصادی به‌کار گرفته شد.

واژگان کلیدی. *داده‌های شمارشی دومتغیره‌ی همبسته؛ مدل پواسون-لگ‌نرمال؛ مدل پواسون-گاما؛ بیز سلسله مراتبی.*

فصل ۱

کلیات پژوهش

۱.۱ مقدمه

تاکنون روش‌های متعددی برای تحلیل داده‌های پیوسته مورد استفاده قرار گرفته است حال آن‌که برای داده‌های گسسته، روش‌های محدودی وجود دارد. لازم به ذکر است که مدل‌های مطرح در تحلیل این گونه داده‌ها نیز در سال‌های اخیر مورد بحث و بررسی قرار گرفته‌اند. یکی از اساسی‌ترین این مدل‌ها ، مدل رگرسیون پواسون می‌باشد که در فصل دوم این مدل به صورت اجمالی مطرح شده و در فصل سوم پارامترهای مدل پیش‌نهادی با استفاده از روش بیز سلسله مراتبی براورد می‌شود و در فصل چهارم مدل رگرسیون پواسون مورد نظر در داده‌های واقعی مربوط به داده‌های اقتصادی بخش تعاون مورد استفاده قرار گرفته و نتایج حاصل از آن بیان خواهد شد و در انتها نتیجه‌ها و پیش‌نهادهای حاصل از این کار ارایه خواهد شد.

۲.۱ مدل رگرسیون پواسون

مدل رگرسیون پواسون در واقع ساده‌ترین و اساسی‌ترین مدل برای تحلیل داده‌های شمارشی است. به عبارت دیگر، اگر داده‌های شمارشی از یک مکانیزم تصادفی تولید شده باشند آن‌گاه انتظار می‌رود که توزیع پواسون برای مدل‌سازی این‌گونه داده‌ها مناسب باشد. در واقع این مدل به عنوان نمونه‌ای از مدل‌های تعمیم یافته‌ی خطی است که شرط اصلی و اساسی استفاده از آن، برابر بودن میانگین و واریانس متغیر وابسته می‌باشد. در این مدل لگاریتم نرخ توزیع پواسون بر اساس مجموع وزنی متغیرهای مستقل مدل‌سازی می‌شود.کریستاسن(۱۹۹۲) برای اولین بار مدل رگرسیون پواسون را برای داده‌های از کارافتادگی دستگاه‌ها به کارگرفت. در سال‌های بعد، مدل رگرسیون پواسون تک متغیره توسط وینکلمان (۲۰۰۳) و در بسیاری از جزئیات دیگر آن توسط گرین (۲۰۰۸) مورد بررسی قرار گرفته است.

همان‌طور که گفته شد برابر بودن میانگین و واریانس متغیر وابسته در این مدل، شرط اصلی استفاده از آن می‌باشد حال آن‌که برقرار نبودن این شرط منجر به بیش پراکندگی در مدل می‌شود. عموماً برای غلبه بر

این مشکل استفاده از مدل‌های آمیخته با استفاده از روش‌های متغیر پنهان پیشنهاد شده است. از جمله این روش‌ها استفاده از مدل دو جمله‌ای منفی (NB) است که به صورت استاندارد برای تحلیل داده‌های شمارشی مورد استفاده قرار می‌گیرد. (میائو، ۱۹۹۴ و وینکلمان، ۲۰۰۰)

راه‌های براورد پارامترهای مدل (NB) و استنباط در مورد آن در مقاله‌های گرین (۲۰۰۷)، هایب (۲۰۰۷) و کمرون و تریودی (۲۰۰۵) مورد بحث قرار گرفته است.

یکی دیگر از روش‌های رایج برای غلبه بر مشکل بیش‌پراکندگی، استفاده از توزیع‌های آمیخته‌ی پواسون - گاما و پواسون -لگ نرمال می‌باشد. در این دو مدل توزیع داده‌ها پواسون در نظر گرفته می‌شود و به این ترتیب که اولا متغیرهای پنهان مستقل در نظر گرفته می‌شوند و ثانیاً به صورت اثرهای تصادفی وارد مدل می‌شوند به طوری‌که داریم:

$$Y_i|\mu_i, h_i \sim P(\mu_i h_i)$$

$$h_i|\cdot \sim D(\cdot)$$

کاس و هسلاگر (۱۹۹۵) از میان توزیع‌های مختلف $D(\cdot)$ در توزیع گاما و لگ نرمال را برای h_i پیش‌نهاد کردند. در این دو مدل به راحتی ثابت می‌شود که توزیع حاشیه‌ای دو جمله‌ای منفی برابر با توزیع پواسون - گاما و پواسون - لگ نرمال بوده و نیز با در نظر گرفتن این توزیع‌های آمیخته بیش پراکندگی در مدل تحت کنترل قرار خواهد گرفت.(اثبات در فصل دوم آورده شده است.)

مسئله‌ی تشخیص بین توزیع‌های مختلف برای متغیرهای پیوسته در مقاله‌های متعددی مورد بحث قرار گرفته است (مارشال و همکاران، ۲۰۰۱ و کندو وهمکاران ۲۰۰۵). اما برای متغیرهای گسسته تحقیق اندک می‌باشد. نیکولولوپولوس و کارلیس (۲۰۰۷) به بررسی توزیع‌های گسسته مشهور پرداخته و آن‌ها را با یکدیگر مقایسه کرده‌اند. از دید آن‌ها هنگامی که با داده‌های شمارشی کار می‌شود به نظر می‌آید که بیش از یک

مدل مناسب برای برازش به داده‌ها وجود دارد، بنابر این تشخیص و تفاوت بین مدل‌های مختلف حساس و دقیق به نظر می‌رسد. برای مدل‌بندی آماری داده‌های پیوسته‌ی چندمتغیره توزیع نرمال نقش یک توزیع پایه‌ای را ایفا می‌کند در صورتی که برای داده‌های شمارشی چندمتغیره، توزیع گسسته چندمتغیره خاصی وجود ندارد که بتواند نقش توزیع پواسون تک‌متغیره را در داده‌های چندمتغیره ایفا کند.

توزیع پواسون دو متغیره توسط جانسون و کنز (۱۹۶۹) ارایه گردید که در آن به ترتیب λ_1 نرخ متغیر اول، λ_2 نرخ متغیر دوم و λ_3 همبستگی بین دو متغیر اول و دوم را بیان می‌کند. همین‌طورتوزیع‌های پواسون دو متغیره‌ای توسط کوچرلاکوتا و کوچرلاکوتا (۱۹۹۲) ارائه گردید. شایان ذکر است در تمامی توزیع‌های ارایه شده ضریب همبستگی محدود به اعداد مثبت می‌باشد.

اگر بر اساس اطلاعات پیشین دریابیم که فرض داشتن رابطه‌ی متقابل بین داده‌های گسسته موجود در مدل پذیرفتنی است، در این صورت استفاده از مدل رگرسیون چندمتغیره‌ی شمارشی مناسب به نظر می‌رسد. راه‌های متعددی برای معرفی این همبستگی‌ها در مدل پواسون چندمتغیره وجود دارد. همان‌طور که گفته شد در مطالعه‌های صورت گرفته‌ی گذشته ضریب همبستگی بین y_1 و y_2 (متغیر پاسخ اول و دوم) در مدل پواسون دو متغیره ،محدود به اعداد مثبت بوده است. کارلیس (۲۰۰۳) مدل پواسون آمیخته‌ی دو متغیره‌ای را ارایه و نشان داد که با استفاده از این روش علاوه بر کنترل بیش پراکندگی مدل، همبستگی بین دو متغیر y_1 و y_2 در مدل آمیخته‌ی حاصل تنها محدود به اعداد مثبت نبوده و همبستگی‌های منفی را نیز پوشش می‌دهد. روش‌های مختلفی برای براورد پارامترهای مدل رگرسیون پواسون وجود دارد. از جمله هو و سینگر (۲۰۰۱) و کوچرلاکوتا و کوچرلاکوتا(۲۰۰۱) به ترتیب روش نیوتن - رافسون و کم‌ترین توان‌های دوم تعمیم یافته را برای ماکسیمم کردن تابع درست‌نمایی به کار بردند. جانگ و وینکلمان (۱۹۹۳) مدل رگرسیون پواسون دو متغیره‌ای را معرفی و طراحی کردند که از روش نیوتن - رافسون برای براورد استفاده می‌کرد. کارلیس

(۲۰۰۳) الگوریتم EM را برای حل مسئله همگرایی که در روش نیوتن - رافسون ایجاد می‌شود، به کار بردند.

چیب و همکاران (۲۰۰۱)، تسیوناس (۲۰۰۵) و کارلیس و انتی زوفراس (۲۰۰۵) استنباط بیزی را برای براورد پارامترهای مدل پیش‌نهاد کردند. در دهه‌های اخیر تحلیل بیزی به دلیل داشتن بعضی از پیش فرض‌های غیر ضروری تحلیل کلاسیک و کارایی این تحلیل برای اندازه‌ی نمونه‌ای کوچک مورد توجه محققین قرار گرفته است. علاوه بر این، روش استنباط بیزی با استفاده از روش مونت کارلوی زنجیر مارکوفی(MCMC) می‌تواند برای ابعاد زیادی از داده‌های شمارشی همبسته نیز به کار گرفته شود. استفاده از این روش جای‌گزین روش‌های براورد کلاسیک شده است.

ماوککلمن (۲۰۰۶) در قالب شبیه‌سازی MCMC با روش نمونه‌گیری گیبزی و الگوریتم متروپولیس - هستینگز را را برای براورد پارامترهای مدل پیشنهادی خود به کار بردند.

خفری، کاظم نژاد و اسکندری (۲۰۰۸) مدل بیزی سلسله مراتبی را برای مدل رگرسیون آمیخته‌ی پواسون - لگ نرمال برای داده‌های شمارشی همبسته به کار بردند. درفصل سوم و چهارم نیز ما از روش MCMC همچنین با بهره‌گیری از روش بیز سلسله مراتبی به براورد پارامترهای مدل مورد نظر خواهیم پرداخت.

۳.۱ هدف‌های پژوهش

هدف از انجام این تحقیق ارایه‌ی یک مدل آماری مناسب با در نظر گرفتن همبستگی دو متغیر تصادفی شمارشی است، که این همبستگی به صورت متغیر تصادفی پنهان یا مولفه‌ی ناهمگنی غیر مشاهده وارد مدل می‌شود. این موضوع به طور خاص برای توزیع‌های دو متغیره‌ی پواسون و در حالتی که از یک ساختار مدل خطی تعمیم یافته برای بررسی میزان اثرگذاری عامل‌های کمکی بر متغیر پاسخ استفاده شده است. در این پژوهش به دلیل کاربرد رهیافت بیزی برای حجم نمونه‌ی کوچک و نیز دیگر مزایای این روش برای براورد