

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

«فرم تأییدیه اعضای هیأت داوران مندرج در پایان نامه کارشناسی ارشد»

بدینوسیله پایان نامه کارشناسی ارشد خانم ندا گیلانی رشته: آمار زیستی گرایش: ----- تقدیم می شود. اینجانبان نسخه نهائی این پایان نامه را از نظر فرم و محتوی بررسی و تأیید کرده و پذیرش آنرا برای تکمیل درجه کارشناسی ارشد پیشنهاد می کنیم.

نام و نام خانوادگی و امضاء اعضای هیأت داوران:

دکتر انوشیروان کاظم نژاد (استاد راهنما)

دکتر فرید زائری (استاد مشاور)

دکتر حمید علوی مجد (استاد ناظر)

دکتر سقراط فقیه زاده (استاد ناظر و نماینده تحصیلات تکمیلی)

## آیین نامه حق مالکیت مادی و معنوی در مورد نتایج پژوهشهای علمی

### دانشگاه تربیت مدرس

**مقدمه:** با عنایت به سیاست‌های پژوهشی و فناوری دانشگاه در راستای تحقق عدالت و کرامت انسانها که لازمه شکوفایی علمی و فنی است و رعایت حقوق مادی و معنوی دانشگاه و پژوهشگران، لازم است اعضای هیأت علمی، دانشجویان، دانش‌آموختگان و دیگر همکاران طرح، در مورد نتایج پژوهشهای علمی که تحت عنوان پایان‌نامه، رساله و طرحهای تحقیقاتی با هماهنگی دانشگاه انجام شده است، موارد زیر را رعایت نمایند:

**ماده ۱-** حق نشر و تکثیر پایان‌نامه/ رساله و درآمدهای حاصل از آنها متعلق به دانشگاه می باشد ولی حقوق معنوی پدید آورندگان محفوظ خواهد بود.

**ماده ۲-** انتشار مقاله یا مقالات مستخرج از پایان‌نامه/ رساله به صورت چاپ در نشریات علمی و یا ارائه در مجامع علمی باید به نام دانشگاه بوده و با تایید استاد راهنمای اصلی، یکی از اساتید راهنما، مشاور و یا دانشجوی مسئول مکاتبات مقاله باشد. ولی مسئولیت علمی مقاله مستخرج از پایان‌نامه و رساله به عهده اساتید راهنما و دانشجو می باشد.

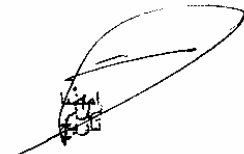
**تبصره:** در مقالاتی که پس از دانش‌آموختگی بصورت ترکیبی از اطلاعات جدید و نتایج حاصل از پایان‌نامه/ رساله نیز منتشر می‌شود نیز باید نام دانشگاه درج شود.

**ماده ۳-** انتشار کتاب و یا نرم افزار و یا آثار ویژه (انری هنری مانند فیلم، عکس، نقاشی و نمایشنامه) حاصل از نتایج پایان‌نامه/ رساله و تمامی طرحهای تحقیقاتی کلیه واحدهای دانشگاه اعم از دانشکده ها، مراکز تحقیقاتی، پژوهشکده ها، پارک علم و فناوری و دیگر واحدها باید با مجوز کتبی صادره از معاونت پژوهشی دانشگاه و براساس آئین‌نامه- های مصوب انجام شود.

**ماده ۴-** ثبت اختراع و تدوین دانش فنی و یا ارائه یافته ها در جشنواره‌های ملی، منطقه‌ای و بین‌المللی که حاصل نتایج مستخرج از پایان‌نامه/ رساله و تمامی طرح‌های تحقیقاتی دانشگاه باید با هماهنگی استاد راهنما یا مجری طرح از طریق معاونت پژوهشی دانشگاه انجام گیرد.

**ماده ۵-** این آیین‌نامه در ۵ ماده و یک تبصره در تاریخ ۸۷/۴/۱ در شورای پژوهشی و در تاریخ ۸۷/۴/۲۳ در هیأت رئیسه دانشگاه به تایید رسید و در جلسه مورخ ۸۷/۷/۱۵ شورای دانشگاه به تصویب رسیده و از تاریخ تصویب در شورای دانشگاه لازم‌الاجرا است.

«اینجانب ندا گیلانی دانشجوی رشته آمار زیستی ورودی سال تحصیلی ۱۳۸۶ مقطع کارشناسی ارشد دانشکده علوم پزشکی متعهد می شوم کلیه نکات مندرج در آیین نامه حق مالکیت مادی و معنوی در مورد نتایج پژوهش های علمی دانشگاه تربیت مدرس را در انتشار یافته های علمی مستخرج از پایان نامه / رساله تحصیلی خود رعایت نمایم. در صورت تخلف از مفاد آیین نامه فوق الاشعار به دانشگاه وکالت و نمایندگی می دهم که از طرف اینجانب نسبت به لغو امتیاز اختراع بنام بنده و یا هرگونه امتیاز دیگر و تغییر آن به نام دانشگاه اقدام نماید. ضمناً نسبت به جبران فوری ضرر و زیان حاصله براساس برآورد دانشگاه اقدام خواهم نمود و بدینوسیله حق هرگونه اعتراض را از خود سلب نمودم.»

  
امضاء  
تاریخ  
۸۹۱۲۲۰

## آیین نامه پایان نامه (رساله) های دانشجویان دانشگاه تربیت مدرس

نظر به اینکه چاپ و انتشار پایان نامه (رساله) های تحصیلی دانشجویان دانشگاه تربیت مدرس، مبین بخشی از فعالیت های علمی پژوهشی دانشگاه است. بنابراین به منظور آگاهی و رعایت حقوق دانشگاه، دانش آموختگان این دانشگاه نسبت به رعایت موارد ذیل متعهد می شوند:

ماده ۱: در صورت اقدام به چاپ پایان نامه (رساله) ای خود، مراتب را قبلاً به طور کتبی به دفتر "دفتر نشر آثار علمی" دانشگاه اطلاع دهد.

ماده ۲: در صفحه سوم کتاب (پس از برگ شناسنامه)، عبارت ذیل را چاپ کند:  
"کتاب حاضر، حاصل پایان نامه کارشناسی ارشد نگارنده در رشته آمار زیستی است که در سال ۱۳۸۹ در دانشکده علوم پزشکی دانشگاه تربیت مدرس به راهنمایی دکتر انوشیروان کاظم نژاد، مشاوره دکتر فرید زایری از آن دفاع شده است.

ماده ۳: به منظور جبران بخشی از هزینه های انتشارات دانشگاه، تعداد یک درصد شمارگان کتاب (در هر نوبت چاپ) را به "دفتر نشر آثار علمی" دانشگاه اهداء کند. دانشگاه می تواند مازاد نیاز خود را به نفع مرکز نشر در معرض فروش قرار دهد.

ماده ۴: در صورت عدم رعایت ماده ۳، ۵۰٪ بهای شمارگان چاپ شده را به عنوان خسارت به دانشگاه تربیت مدرس، تادیه کند.

ماده ۵: دانشجو تعهد و قبول می کند در صورت خودداری از پرداخت های بهای خسارت، دانشگاه مذکور را از طریق مراجع قضایی مطالبه و وصول کند، به علاوه به دانشگاه حق می دهد به منظور استیفای حقوق خود، از طریق دادگاه، معادل وجه مذکور در ماده ۴ را از محل توقیف کتابهای عرضه شده نگارنده برای فروش، تامین نماید.

ماده ۶: اینجانب ندا گیلانی دانشجوی رشته آمار زیستی مقطع کارشناسی ارشد تعهد فوق و ضمانت اجرایی آن را قبول کرده، به آن ملتزم می شوم.

نام و نام خانوادگی  
تاریخ و امضا  
ندا گیلانی  
۸۹۱۲۰۲۰



دانشگاه تربیت مدرس  
دانشکده علوم پزشکی

پایان نامه دوره کارشناسی ارشد آمار زیستی

عنوان:

استفاده از مدل‌های GMM در تفسیر بیماری پرفشاری خون

نگارش:

ندا گیلانی

استاد راهنما:

دکتر انوشیروان کاظم نژاد

استاد مشاور:

دکتر فرید زایری

بهار ۱۳۸۹

سپاس و ستایش باد بر آفریننده زیبایی ها که ذره ذره عالم ، آیه جمال و جلال اوست.

## تقدیم به بهترین های زندگی ام :

👤 دو گوهر بی مثال زندگی ، مادر فداکار و با گذشت و پدر بزرگوارم که سرمایه

عمر گرانبهایشان را نثارم نمودند .

👤 خواهر عزیز تر از جانم که همواره دوستش دارم .

## تشکر و قدردانی

خداوند، سستی آفرین راسپاس  
که به من ارزانی داشت  
خوشه چینی و دانش پژوهی از محضر استاید گرانقدر

زیباترین گل واژه های قدرشناسی و سپاسی خود را با نهایت فروتنی تقدیم می کنم به:

- ❖ استاد ارجمند و اسوه اخلاق، جناب آقای دکتر نوشیروان کاظم نژاد به خاطر راهنمایی های کارساز و دلسوزی های همه جانبه ای که در حق بنده روا داشتند، همچنین استاد محترم جناب آقای دکتر فرید زایری که با نظرات ارزنده و حمایت های خویش، راهگشای اینجانب بوده اند.
  - ❖ استاید ارجمند جناب آقای دکتر سطرط هفید زاده و دکتر ابراهیم حاجی زاده به خاطر حسن توجه ها و محبت هایی که نسبت به بنده و سایر دانشجویان در دوران تحصیل روا داشته اند.
  - ❖ تمامی کسانی که مراد نخواست این پایان نامه یاری کرده اند؛ علی الخصوص جناب آقای دکتر جمشید زندانی که داده های مربوط به طرح بررسی عوامل خطر بیماری های غیر واکسیروار را در اختیارم قرار نهادند.
- و سرانجام بر خود فرض می دانم از پدر و مادر و ماد خدا کلام، خواهر و داماد مهربانان که با صبر و حوصله و بزرگواری سرشارشان، توفیق کسب این درجه علمی را نصیب بنده نمودند، بطور ویژه تشکر و قدردانی نمایم.

نذاکیلانی

بهار ۸۹

## چکیده

مدل‌ها نمایش ساده و خلاصه‌ای از واقعیت هستند که در اغلب علوم و تکنولوژی کاربرد دارند. هدف از مدل‌سازی معمولاً توصیف اثرهای متغیرهای توضیحی روی متغیرهای پاسخ است. یکی از مهمترین آنها، مدل‌های آمیخته خطی تعمیم یافته است. مدل‌های آمیخته خطی تعمیم یافته کاربرد وسیعی در تحلیل داده‌های خوشه‌بندی شده شامل داده‌های طولی یا اندازه‌گیری مکرر دارند. این مدل برای مواردی با پاسخ‌های غیر نرمال و تابع پیوند غیر خطی میان میانگین پاسخ و متغیرهای توضیحی، محاسبه بیش پراکنشی و همبستگی ناشی از وجود اثرهای تصادفی، مفید بنظر می‌رسد.

با استفاده از زمینه نظری فوق به بررسی عوامل موثر بر بیماری پرفشاری خون - که در بردارنده پاسخ‌های از نوع دو حالتی هستند - با استفاده از برازش دو مدل لجستیک با اندازه‌گیری مکرر (حاشیه ای) و مدل لجستیک با اثر تصادفی (شرطی) مبادرت شده است. یافته‌های این تحقیق با استفاده از معیارهای نیکویی برازش نشان می‌دهد که برازش مدل شرطی تا حدودی بهتر از مدل حاشیه‌ای است. متغیرهای جنسیت، گروه سنی، قند خون ناشتا فعالیت بدنی، نمایه توده بدنی و مصرف میوه-سبزی در مدل لجستیک حاشیه ای در سطح ۵ درصد معنی‌دار آماری شدند. در مدل لجستیک شرطی با روش برآورد شبه درست‌نمایی مقید تنها دو متغیر گروه سنی و نمایه توده بدنی در این سطح از لحاظ آماری معنی‌دار گردیدند و در مدل لجستیک شرطی با استفاده از تقریب لاپلاس چهار متغیر گروه سنی، فعالیت بدنی، نمایه توده بدنی و مصرف میوه-سبزی به عنوان عوامل مرتبط با بیماری پرفشاری خون شناخته شدند و دیگر متغیرهای زمینه‌ای مورد بررسی از لحاظ آماری معنی‌داری نشان ندادند.

**کلمات کلیدی:** مدل‌های آمیخته خطی تعمیم‌یافته، معیارهای برازش، مدل لجستیک، بیماری پرفشاری

خون، فشار خون.



## فهرست مطالب

### عنوان

- فصل اول : مقدمه و مروری بر مطالعات انجام شده.....۱
- ۱-۱. کلیات مفاهیم آماری.....۲
- ۱-۱-۱. مقدمه.....۲
- ۱-۱-۲. مدل‌های خطی تعمیم یافته.....۲
- ۱-۱-۳. مدل‌های آمیخته خطی تعمیم یافته.....۲
- ۱-۱-۴. داده‌های خوشه‌ای در مقابل داده‌های با اندازه‌گیری مکرر و طولی.....۳
- ۱-۱-۵. سطوح داده‌ای.....۴
- ۱-۱-۶. انواع عامل‌ها و اثرهای متناظر با آنها.....۴
- ۱-۲. کلیات پزشکی.....۵
- ۱-۲-۱. آشنایی با فشار خون.....۵
- ۱-۲-۲. شرح بیماری.....۷
- ۱-۲-۳. عوامل عمومی.....۸
- ۱-۲-۴. عوامل مستعد کننده بیماری.....۹
- ۱-۲-۵. علائم و عوارض بیماری.....۹
- ۱-۳. فرضیه تحقیق.....۱۰
- ۱-۴. سوالات تحقیق.....۱۰
- ۱-۵. اهداف تحقیق.....۱۱
- ۱-۶. چارچوب مطالعه.....۱۱
- ۱-۶-۱. نوع مطالعه.....۱۱

۱۱	۲-۶-۱. جامعه آماری
۱۱	۳-۶-۱. واحد آماری
۱۲	۷-۱. روش گردآوری اطلاعات
۱۲	۱-۷-۱. حجم نمونه
۱۳	۲-۷-۱. روش نمونه گیری
۱۳	۸-۱. متغیرهای تحقیق
۱۶	۹-۱. تاریخچه مختصر از GLMM ها
۱۶	۱-۹-۱. پیشرفت‌های مهم از لحاظ تئوری
۱۸	۲-۹-۱. پیشرفت‌های مهم از لحاظ نرم افزاری
۱۸	۱۰-۱. بررسی اطلاعات موجود و پژوهش‌های انجام شده در زمینه فشار خون
۱۸	۱-۱۰-۱. مطالعات انجام یافته در ایران
۲۳	۲-۱۰-۱. پژوهش‌های انجام شده در سایر کشورها
۲۷	<b>فصل دوم: چارچوب نظری تحقیق</b>
۲۸	۱-۲. مقدمه
۲۸	۲-۲. مدل‌های خطی تعمیم یافته
۲۹	۱-۲-۲. مؤلفه‌های یک مدل خطی تعمیم یافته
۳۰	۲-۲-۲. توابع پیوند
۳۱	۳-۲-۲. تابع لگ درستنمایی برای GLM ها
۳۱	۴-۲-۲. برازش GLM با روش ML
۳۲	۵-۲-۲. آزمون فرضیه

- ۳۲ ..... ۱-۵-۲-۲. آزمون‌های نسبت درستنمایی
- ۳۲ ..... ۲-۵-۲-۲. آزمون‌های والد
- ۳۳ ..... ۶-۲-۲. فواصل اطمینان
- ۳۳ ..... ۳-۲. مدل‌های آمیخته خطی
- ۳۴ ..... ۴-۲. مدل‌های آمیخته خطی تعمیم یافته
- ۳۵ ..... ۱-۴-۲. ساختار مدل‌های GLMM
- ۳۶ ..... ۲-۴-۲. توزیع شرطی  $y$
- ۳۶ ..... ۳-۴-۲. تابع درستنمایی تحت GLMM
- ۳۷ ..... ۴-۴-۲. استنباط تقریبی
- ۳۷ ..... ۱-۴-۴-۲. تقریب لاپلاس
- ۳۸ ..... ۲-۴-۴-۲. برآورد شبه درستنمایی تاوانیده
- ۴۱ ..... ۳-۴-۴-۲. بیشینه درستنمایی سلسله مراتبی
- ۴۲ ..... ۴-۴-۴-۲. شبه درستنمایی یا درستنمایی نما
- ۴۳ ..... ۵-۴-۲. استنباط بر پایه درستنمایی
- ۴۴ ..... ۱-۵-۴-۲. الگوریتم EM مونت کارلوئی
- ۴۵ ..... ۲-۵-۴-۲. بیشینه سازی با روش جزءبه‌جزء
- ۴۷ ..... ۶-۴-۲. پیش‌بینی اثرات تصادفی
- ۴۶ ..... ۱-۶-۴-۲. برآورد توأم اثرهای ثابت و تصادفی
- ۴۸ ..... ۷-۴-۲. استنباط به روش بیزی
- ۵۰ ..... ۱-۷-۴-۲. برآورد پارامتر، فواصل احتمالاتی و  $p$ -مقدارها
- ۵۱ ..... ۲-۷-۴-۲. تعیین توزیع پیشین‌های بی اطلاع
- ۵۶ ..... ۳-۷-۴-۲. تعیین توزیع پسین

۵۷	.....۴-۷-۴-۲. استنباط بیزی در GLMM ها
۵۸	.....۸-۴-۲. معادلات برآورد
۵۹	.....۱-۸-۴-۲. معادلات برآورد تعمیم یافته
۶۲	.....۲-۸-۴-۲. معادلات برآورد تکراری
۶۳	.....۳-۸-۴-۲. روش گشتاورهای شبیه سازی شده
۶۳	.....۴-۸-۴-۲. برآورد استوار در GLMM
۶۳	.....۵-۸-۴-۲. ضرایب همبستگی
۶۴	.....۱-۵-۸-۴-۲. انواع همبستگی
۶۴	.....۹-۴-۲. انتخاب GLMM
۶۴	.....۱-۹-۴-۲. اصل عمومی برای انتخاب مدل
۶۷	.....۱۰-۴-۲. مؤلفه‌های واریانس منفی
۶۸	..... فصل سوم: نتایج و یافته‌ها
۶۹	.....۱-۳. مقدمه
۶۹	.....۲-۳. مشخصات مطالعه
۷۱	.....۳-۳. روش تجزیه و تحلیل اطلاعات
۷۳	.....۱-۳-۳. نتایج مدل لجستیک با معادلات برآورد تعمیم یافته
۷۶	.....۱-۱-۳-۳. برازش مدل لجستیک با اندازه‌گیری مکرر
۷۷	.....۲-۱-۳-۳. نمودارهای مانده‌ای حاشیه‌ای مدل لجستیک با اندازه‌گیری مکرر
۷۸	.....۲-۳-۳. مدل لجستیک شرطی با اثر تصادفی
۷۸	.....۱-۲-۳-۳. نتایج مدل لجستیک شرطی به روش PL
۸۰	.....۲-۲-۳-۳. نتایج مربوط به مدل لجستیک شرطی با روش تقریب لاپلاس
۸۳	.....۳-۲-۳-۳. بررسی برازش نتایج مدل لجستیک با اثر تصادفی

۸۴	.....۴-۲-۳-۳. نمودارهای مانده‌ای شرطی مدل لجستیک با خوشه‌های تصادفی
۸۵	.....۳-۳. مقایسه مدل‌های آماری در تعیین عوامل مرتبط با بیماری فشار خون بالا
۹۰	..... فصل چهارم: بحث، نتیجه‌گیری و پیشنهادها
۹۱	.....۱-۴. پاسخ به سوالات پژوهش
۹۱	.....۱-۴-۱. مقایسه مدل‌های آماری در تعیین عوامل مرتبط با پرفشاری خون
۹۳	.....۲-۴. مقایسه نتایج حاصله با پژوهش‌های قبلی
۹۶	.....۳-۴. پیشنهادها
۹۷	..... فهرست منابع
۱۰۱	..... ضمائم
۱۱۴	..... چکیده انگلیسی

## فهرست جداول

### عنوان جدول

- جدول ۱-۱. تقسیم بندی فشارخون بر اساس گزارش JNC-7 ..... ۸
- جدول ۱-۲. متغیرهای مورد مطالعه ..... ۱۴
- جدول ۱-۳. نمایی از نمونه به تفکیک شهرستان، خوشه، گروههای سنی و جنسی ..... ۷۰
- جدول ۲-۳. توزیع فراوانی افراد بر اساس داشتن بیماری فشارخون و متغیرهای پایه رسته‌ای ..... ۷۲
- جدول ۳-۳. نتیجه برازش مدل لجستیک اندازه‌گیری مکرر با استفاده از روش Residual PL ..... ۷۴
- جدول ۳-۴. Type III Testes of fixed effects ..... ۷۵
- جدول ۳-۵. آماره نیکویی برازش ..... ۷۶
- جدول ۳-۶. نتایج برازش مدل لجستیک با اثر تصادفی افراد با استفاده از روش Residual PL ..... ۷۹
- جدول ۳-۷. Type III Testes of fixed effects برای مدل لجستیک با اثر تصادفی افراد با استفاده از روش Residual PL ..... ۸۰
- جدول ۳-۸. نتیجه برازش مدل لجستیک با اثر تصادفی افراد با استفاده از روش لاپلاس ..... ۸۱
- جدول ۳-۹. Type III Testes of fixed effects برای مدل لجستیک با اثر تصادفی افراد با استفاده از روش لاپلاس ..... ۸۲
- جدول ۳-۱۰. آماره‌های نیکویی برازش ..... ۸۳
- جدول ۳-۱۱. نتایج سه مدل برای بررسی عوامل مرتبط با بیماری پر فشارخون ..... ۸۶
- جدول ۳-۱۲. معیارهای اطلاع ..... ۸۷
- جدول ۳-۱۳. مقایسه مدل‌ها با استفاده از آماره‌های نیکویی برازش ..... ۸۸
- جدول ۳-۱۴. معیارهای برازش GEE ..... ۸۹

## فهرست نمودارها

### عنوان

- نمودار ۲-۱. درستنمایی در مقابل میانگین..... ۵۳
- نمودار ۲-۲. درستنمایی در مقابل سیگما..... ۵۴
- نمودار ۲-۳. درستنمایی در مقابل لگ سیگما..... ۵۴
- نمودار ۳-۱. نمودار مانده‌ای استاندارد شده حاشیه‌ای مدل لجستیک با اندازه‌گیری مکرر..... ۷۷
- نمودار ۳-۲. نمودار مانده‌ای پیرسون حاشیه‌ای مدل لجستیک با اندازه‌گیری مکرر..... ۷۷
- نمودار ۳-۳. نمودار مانده‌ای استاندارد شده شرطی مدل لجستیک با خوشه‌های تصادفی..... ۸۴
- نمودار ۳-۴. نمودار مانده‌ای پیرسون شرطی مدل لجستیک با خوشه‌های تصادفی..... ۸۵



فصل اول

مقدمه و مروری  
بر مطالعات انجام شده



## ۱-۱. کلیات مفاهیم آماری

### ۱-۱-۱. مقدمه

مدل‌ها نمایش ساده و خلاصه‌ای از واقعیت هستند که در اغلب علوم و تکنولوژی کاربرد دارند. اگرچه اطمینان کامل به درستی یک مدل، تصویری نابجاست؛ با این وجود بیشتر استنباط‌های آماری بر اساس این فرض شکل می‌گیرند. مدل‌های با مؤلفه احتمالاتی را مدل‌های آماری می‌نامند. هدف از مدل‌سازی معمولاً توصیف اثرهای متغیرهای توضیحی روی متغیرهای پاسخ است. یکی از مهمترین آن‌ها، مدل‌های خطی تعمیم‌یافته است.

### ۱-۱-۲. مدل‌های خطی تعمیم یافته<sup>۱</sup> (GLM)

خانواده مدل‌های خطی تعمیم یافته، اولین بار توسط Nelder and Wedderburn (1972) معرفی شدند. این خانواده، نشان دهنده رده‌ای از مدل‌های رگرسیونی با اثرات ثابت برای انواعی از متغیرهای وابسته (پیوسته، دوحالتی و شمارشی) هستند؛ که مدل‌های رگرسیون خطی با خطای نرمال، مدل‌های رگرسیونی نمایی غیر خطی، لجستیک، پواسن و بسیاری از مدل‌های دیگر مانند مدل‌های لگ خطی برای داده‌های رسته‌ای را در بر می‌گیرد [۱].

### ۱-۱-۳. مدل‌های آمیخته خطی تعمیم یافته<sup>۲</sup> (GLMM)

از این مدل‌ها زمانی استفاده می‌شود که بخواهیم اثرهای تصادفی را بعنوان بخشی از پیش بین خطی در مدل‌های خطی تعمیم یافته قرار دهیم. GLMM را می‌توان با تعیین چهار مورد زیر تبیین نمود:

<sup>۱</sup> - Generalized Linear Models

<sup>۲</sup> - Generalized Linear Mixed Models

۱. تعیین توزیع داده ای برای مقادیر ثابت متغیرهای پیشگو

۲. تعیین اینکه چه جنبه‌ای از مسئله مدل‌بندی شود.

۳. تعیین متغیرهای پیشگوی مدل

۴. تعیین متغیرهای پیشگوی رسته‌ای که دارای توزیع تصادفی هستند.

جزئیات بیشتر و نحوه استنباط این مدل‌ها را در فصل ۲ بیان خواهیم کرد [۲].

## ۱-۴. داده‌های خوشه‌ای در مقابل داده‌های با اندازه‌گیری مکرر و طولی

اصطلاح داده‌های خوشه‌ای به مجموعه داده‌ای اطلاق می‌شود که در آن متغیر وابسته برای هر آزمودنی برای یکبار اندازه‌گیری می‌شود و واحدهای تحلیل بطور گروه‌بندی یا آشیانی درون خوشه‌ها قرار می‌گیرند. بعنوان مثال نمرات ریاضی دانشجویان (واحدهای تحلیل) در کلاس‌های درسی (واحدهای خوشه‌ای) آشیانی می‌شوند؛ که کلاس‌ها نیز به نوبه خود در مدارس (خوشه‌های خوشه‌ها) آشیانی شده‌اند که این نوع داده را مجموعه داده خوشه‌ای سه سطحی می‌نامند.

داده‌های با اندازه‌گیری مکرر، عموماً داده‌هایی هستند که متغیر وابسته برای هر آزمودنی بیش از یکبار در میان سطوح یک عامل (یا عوامل) با اندازه‌های تکراری سنجیده می‌شود. عامل‌های با اندازه مکرر، زمان یا سایر شرایط آزمایشی یا مشاهده‌ای هستند که اغلب عوامل درون آزمودنی<sup>۱</sup> نامیده می‌شوند.

داده‌های طولی در واقع مجموعه داده‌هایی هستند که در آن‌ها متغیر وابسته در چند مرحله زمانی برای هر واحد تحلیل اندازه گرفته می‌شود. معمولاً این نوع داده‌ها بصورت حداقل دو اندازه‌گیری مکرر در طول بازه زمانی طولانی روی می‌دهند.

---

<sup>۱</sup> - Within-subject factors

داده‌های طولی خوشه‌بندی شده ویژگی‌های داده‌های خوشه‌ای و طولی را با هم ترکیب می‌کند. خصوصاً اینکه واحدهای تحلیل درون خوشه‌ها آشیانی می‌شوند و هر واحد بیش از یکبار اندازه‌گیری می‌شود.

داده‌های خوشه‌ای با اندازه‌گیری مکرر و داده‌های طولی را مجموعه داده‌های سلسله مراتبی<sup>۱</sup> می‌نامند؛ چرا که مشاهدات را می‌توان در سطوحی از یک سلسله مراتب داده‌ای قرار داد.

## ۱-۱-۵. سطوح داده‌ای

مفهوم "سطوح" داده‌ای مبتنی بر مدل بندی خطی سلسله مراتبی (Raudenbush and Bryk(2002) است. همه مجموعه‌های داده‌ای حداقل دارای دو سطح برای استنباط هستند. سطح ۱ برای نشان مشاهدات در جزئی‌ترین سطح داده‌ای؛ در داده‌های خوشه‌ای، سطح ۱ نمایانگر اندازه‌گیری‌های مکرری است که روی واحدهای تحلیل یکسان ایجاد شده است. متغیر وابسته پیوسته همواره در سطح ۱ داده‌ای اندازه‌گیری می‌شود. سطح ۲ نشان دهنده سطح بعدی سلسله مراتب است؛ در مجموعه داده‌های خوشه‌ای، سطح ۲ مشاهدات نمایانگر خوشه‌های واحدهاست. سطح ۳، سطح دیگر سلسله مراتب است و معمولاً به خوشه‌های واحدهای سطح ۲ (خوشه‌های خوشه‌ها) در مجموعه داده‌های خوشه‌ای سه سطحی اشاره دارد.

## ۱-۱-۶. انواع عامل‌ها و اثرهای متناظر با آن‌ها

تفاوت بین عامل‌های تصادفی و ثابت حائز اهمیت است. معمولاً از مفهوم عامل ثابت بیشتر در مدل تحلیل واریانس (ANOVA) یا تحلیل مدل کوواریانس (ANCOVA) استفاده می‌شود. عامل ثابت بعنوان یک متغیر رسته‌ای یا رده‌بندی تعریف می‌شود. عامل‌های ثابت ممکن است شامل متغیرهای کمکی کیفی مانند جنسیت، متغیرهای رده بندی در طرح نمونه‌گیری پیمایشی<sup>۲</sup> مانند ناحیه یا طبقه یا متغیرهای رده‌بندی ترتیبی در یک مطالعه مشاهده‌ای مانند گروه سنی باشد.

---

<sup>۱</sup> - hierarchical

<sup>۲</sup> - Survey Sampling

یک عامل تصادفی در واقع یک متغیر رده‌بندی است با سطوحی که می‌توان بعنوان نمونه‌گیری تصادفی از یک جمعیت در نظر گرفته شود. همه سطوح ممکن عامل تصادفی در مجموعه داده‌ای حضور ندارند. متغیرهای رده بندی که واحدهای سطح ۲ و سطح ۳ در داده‌های خوشه‌ای طولی و اندازه‌گیری مکرر را مشخص می‌سازند، معمولاً بعنوان عامل‌های تصادفی در نظر گرفته می‌شوند. اثرهای ثابت که ضرایب رگرسیونی یا پارامترهای اثر ثابت نیز نامیده می‌شوند، رابطه بین متغیر وابسته و متغیرهای پیشگو (عامل‌های ثابت یا متغیرهای کمکی پیوسته) برای تمام یا نسبت کمتری از جامعه تبیین می‌سازند. این اثرها بر اساس داده‌های جمع آوری شده در یک مطالعه تحقیقی، برآورد می‌شوند.

اثرهای تصادفی، مقادیر تصادفی مرتبط با سطوح یک عامل (یا عوامل) تصادفی هستند. برخلاف اثرهای ثابت، اثرهای تصادفی بعنوان متغیرهای تصادفی در مدل قرار می‌گیرند. هرگاه سطح بخصوصی از یک عامل (تصادفی یا ثابت) را بتوان فقط در یک سطح از عامل دیگر اندازه‌گیری کرد، سطوح عامل اول را در اصطلاح آشیانی شده در سطوح عامل دوم می‌نامند. اثرهای عامل آشیانی روی متغیر پاسخ را اثرهای آشیانی می‌نامند. اما اگر سطح بخصوصی از یک عامل (تصادفی یا ثابت) را بتوان میان سطوح مختلف عامل دیگر اندازه‌گیری کرد، اصطلاحاً آن عوامل را متقاطع و اثرهای این عامل‌ها روی متغیر وابسته را اثرهای متقاطع می‌نامند [۳].

## ۲-۱. کلیات پزشکی

### ۱-۲-۱. آشنایی با فشارخون

سرخرگ‌ها خون را از قلب به دیگر نقاط بدن می‌رسانند. برای آن که خون بتواند در این عروق به جریان در آید، نیاز به فشار مناسبی دارد. این فشار جریان خون در سرخرگ‌های بدن یعنی فشاری که در هر انقباض عضله قلب در اثر برخورد خون به دیواره سرخرگ وارد می‌شود، فشارخون نام دارد (شکل ۱-۱). هنگامی که سرخرگ‌های بزرگ، قابلیت ارتجاع و استحکام طبیعی خود را از دست بدهند و عروق کوچک، نیز باریکتر شوند، فشارخون بالا می‌رود. قلب همانند یک پمپ با انقباض و