تاییدیه اعضای هیات داوران حاضر در جلسه دفاع از رساله دکتری

خانم **سمانه حسین زاده** رشته **آمار زیستی** رساله **دکتری** خود را با عنوان « **مدلسازی فواصل زمانی بین حوادث بازگشتی با استفاده از مدل شکنندگی وابسته به زمان و کاربرد آن در داده های پزشکی** » در تاریخ **۱۳۹۱/۸/۱۰** ارائه کردند.

بدینوسیله اعضای هیات داوران نسخه نهایی این رساله را از نظر فرم و محتوا تایید کرده و پذیرش آنرا برای تکمیل درجه دکتری پیشنهاد می‌کنند.

اعضای هیات داوران	نام و نام خانوادگی	امضاء
استاد راهنما	دکتر سقراط فقیه زاده	
استاد مشاور	دکترمهدی رهگذر	
استاد مشاور	دکترابراهیم حاجی زاده	
استاد ناظر	دکترصدیقه السادات طوافیان	
استاد ناظر	دکتریداله محرابی	
استاد ناظر	دکترمسعود صالحی	
استاد ناظر و نماینده تحصیلات تکمیلی	دکترانوشیروان کاظم نژاد	

آیین‌نامه 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 عناوین پایان‌نامه، رساله و طرحهای تحقیقاتی با هماهنگی دانشگاه انجام شده است، موارد زیر را رعایت نمایند:

ماده ۱– حق نشر و تکثیر پایان نامه/ 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 پایان نامه 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 (اثری 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 آئین‌نامه‌های 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 **سمانه حسین زاده** دانشجوی رشته **آمارزیستی** ورودی سال تحصیلی **۱۳۸۶** مقطع **دکتری** 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».

امضا

تاریخ ۹۱/۰۹/۰۷

آئین‌نامه پایان‌نامه (رساله)های دانشجویان دانشگاه تربیت مدرس

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

ماده ۱ : در صورت اقدام به چاپ پایان‌نامه (رساله)ی خود، مراتب را قبلاً به‌طور 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 **رساله دکتری** نگارنده در رشته **آمارزیستی** است که در سال **۱۳۹۱** در دانشکده علوم پزشکی دانشگاه تربیت مدرس به راهنمایی **دکتر سقراط فقیه زاده**، مشاوره **دکتر مهدی رهگذر** و **دکتر ابراهیم حاجی-زاده** از آن دفاع شده است.

ماده ۳ : به منظور جبران بخشی از هزینه‌های انتشارات دانشگاه، تعداد یک درصد شمارگان کتاب (در هر نوبت چاپ) را به "دفتر نشر آثار علمی" دانشگاه اهداء کند. دانشگاه می‌تواند 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 به‌عنوان خسارت به دانشگاه تربیت مدرس، تادیه کند.

ماده ۵ : دانشجو تعهد و قبول می‌کند در صورت خودداری از پرداخت‌های بهای خسارت، دانشگاه مذکور را از طریق مراجع قضایی مطالبه و وصول کند، به علاوه به دانشگاه حق می‌دهد 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 **سمانه حسین زاده** دانشجوی رشته **آمارزیستی** مقطع **دکتری** تعهد فوق و ضمانت اجرایی آن را قبول کرده، به آن ملتزم می‌شوم.

نام و نام خانوادگی: سمانه حسین زاده

تاریخ و امضا: ۹۱/۰۹/۰۷

دانشگاه تربیت مدرس
دانشکده علوم پزشکی

رساله

دوره دکتری تخصصی (Ph.D) در رشته آمارزیستی

عنوان

مدل‌سازی فواصل زمانی بین حوادث بازگشتی با استفاده از مدل شکنندگی وابسته به زمان و کاربرد آن در داده های پزشکی

نگارش

سمانه حسین زاده

استاد راهنما

دکتر سقراط فقیه زاده

اساتید مشاور

دکتر مهدی رهگذر

دکتر ابراهیم حاجی زاده

پاییز ۱۳۹۱

تقدیم و درود فراوان به روح پدر بزرگوارم

تقدیم و سپاس بیکران

بر همدلی و همراهی و پشتیبانی مادر دلسوز و مهربانم

و همسر صبورم که با حمایت های همه جانبه و فراهم آوردن محیطی آرام، مرا در به اتمام

رساندن این مرحله از تحصیل یاری نمود.

و خواهرانم و برادرم که در تمامی دوران زندگی ام امید موفقیت را همواره در من زنده نگاه

داشتند.

بسی شایسته است از استاد فرهیخته و فرزانه جناب آقای دکتر سقراط فقیه زاده که سرای علم و دانش را با راهنمایی های سازنده خویش برای من روشن ساختند؛ تقدیر و تشکر نمایم.

از اساتید فرهیخته جناب آقای دکتر مهدی رهگذر و جناب آقای دکتر ابراهیم حاجی زاده بسیار سپاسگذارم که همواره راهنما و راه گشای اینجانب در اتمام رساله بودند.

و با تقدیر و تشکر شایسته از جناب آقای دکتر سهراب هاشمی فشارکی و سرکار خانم مرضیه قراخانی به دلیل یاری‌ها و راهنمایی‌های بی چشمداشت ایشان.

چکیده

زمینه: در داده‌های حوادث بازگشتی که در آن یک حادثه برای فردی چندین بار تکرار می‌شود، همبستگی به وجود می‌آید. بررسی اثر عوامل خطر بر فاصله زمانی بین حوادث در این داده ها بیشتر مورد توجه است. برای تحقق این هدف استفاده از مدل خطرات متناسب کاکس به خاطر پیش‌فرض استقلال داده‌ها مناسب نیست. با مدل شکنندگی که وابستگی حوادث را در مدل لحاظ می‌کند، استنباط‌های کاراتری بدست می‌آید. یکی از مفروضات این مدل ثابت بودن شکنندگی در طول مدت پیگیری است که می‌تواند ناکافی باشد. این امکان وجود دارد که هر چه فاصله بین حوادث بیشتر شود، وابستگی آنها نیز کمتر شود یا شکنندگی بعد از شکست اول کاهش یا افزایش یابد. لذا مدل‌هایی با شکنندگی وابسته به زمان، واقعی‌تر هستند. هدف این مطالعه بررسی و برازش مدل شکنندگی وابسته به زمان بر فاصله زمانی بین حوادث بازگشتی است.

موادو روش‌ها: حوادث بازگشتی مورد مطالعه ظهور علامت دیسشارژ IED در نوار مغزی ۵۶ بیمار مبتلا به صرع بود. دو مدل شکنندگی وابسته به زمان بر فاصله‌زمانی بین حوادث با روش برآورد مربع‌بندی گاوسی برازش شد. شبیه‌سازی جهت بررسی برآورد مدل‌ها انجام شد.

نتیجه: دو عامل سن و وضعیت جانبازی از عوامل تاثیرگذار بر فاصله‌زمانی بین دیسشارژها هستند. معنی‌دار بودن واریانس شکنندگی نشان می‌دهد که وجود اثر تصادفی در این داده‌ها ضروری است. مدل اول شکنندگی وابسته به زمان مدل مناسبی نبود ولی مدل دوم شکنندگی وابسته به زمان نشان داد که عامل تصادفی با اثر جمعی یک عامل اثرگذار بر فاصله زمانی بین حوادث است ولی شکنندگی با اثر ضربی معنی‌دار نبود. شبیه سازی اریبی کمتری را برای مدل دوم شکنندگی وابسته به زمان نشان داد.

بحث: مدل شکنندگی وابسته به زمان بهتر از مدل شکنندگی برازش شد. روش برآورد کوادراتور گاوسی یک تکنیک مناسب و کاربردی برای برازش مدل‌های شکنندگی وابسته به زمان است.

کلمات کلیدی: شکنندگی وابسته به زمان، حوادث بازگشتی، کوادراتور گاوسی، صرع

فهرست مطالب

فهرست ..... صفحه

فصل ۱ مقدمه و کلیات ۱

۱-۱ مقدمه ۲

۱-۲ بیان مساله و ضرورت اجرا ۵

۱-۳ مساله پژوهش از دیدگاه پزشکی (صرع) ۹

۱-۴ روش انجام تحقیق ۱۲

۱-۵ روش و ابزار گردآوری اطلاعات ۱۲

۱-۶ محاسبات آماری ۱۳

۱-۷ اهداف تحقیق ۱۳

۱-۸ فرضیه‌ها ۱۳

فصل ۲ مدل‌های حوادث بازگشتی و پیشینه تحقیق ۱۴

۲-۱ مقدمه ۱۵

۲-۲ انواع مدل ها ۲۱

۲-۲-۱ معرفی نمادها ۲۱

۲-۲-۲ مدل فرآیند شمارشی ۲۵

۲-۲-۳ مدل‌های تصحیح واریانس ۲۷

۲-۲-۴ شکنندگی و انواع مدل‌های شکنندگی ۳۲

۲-۳ انتخاب مدل مناسب ۴۱

۲-۴ مدل برای فاصله زمانی ۴۳

۲-۵ روش برآورد ۴۵

۲-۶ تابع خطر پایه ۵۳

۲-۷ پیشیینه تحقیق ۵۵

فصل ۳ آنالیز عددی و مدل شکنندگی وابسته به زمان ۶۷

۳-۱ مقدمه آنالیز عددی ۶۸

۳-۲ انتگرال‌گیری عددی ۶۸

۳-۲-۱ چندجمله‌ای درونیاب ۶۹

۳-۲-۲ چندجمله‌ای‌های متعامد ۶۹

۳-۲-۳ انواع روش های انتگرال گیری عددی ۷۰

۳-۲-۴ طرح‌های سازگار ۷۶

۳-۲-۵ خطا در محاسبه تقریب انتگرال ۷۸

۳-۳ بهینه‌سازی ۷۹
۳-۳-۱ تعریف اصطلاحات ۸۰
۳-۳-۲ انواع روش‌های بهینه سازی ۸۲
۳-۳-۳ معیار همگرایی ۹۷
۳-۳-۴ سرعت یا مرتبه همگرایی ۹۹
۳-۳-۵ مقدار آغازین یا اولیه ۹۹
۳-۳-۶ کاربرد روشهای بهینه سازی در آمار ۱۰۰
۳-۴ مدل شکنندگی وابسته به زمان ۱۰۵
۳-۴-۱ مدل اول شکنندگی وابسته به زمان ۱۰۶
۳-۴-۲ مدل دوم شکنندگی وابسته به زمان ۱۱۳

فصل ۴ یافته ها ۱۱۹
۴-۱ نمونه‌های بیماران مبتلا به صرع ۱۲۰
۴-۱-۱ توصیف داده‌ها ۱۲۲
۴-۱-۲ تحلیل داده‌ها ۱۳۰
۴-۲ شبیه سازی ۱۴۷
۴-۲-۱ شبیه سازی در مدل اول شکنندگی وابسته به زمان ۱۴۹
۴-۲-۲ شبیه سازی در مدل دوم شکنندگی وابسته به زمان ۱۵۲

فصل ۵ بحث و نتیجه گیری ۱۵۴
۵-۱ بحث ۱۵۵
۵-۲ پیشنهادات ۱۶۴

فهرست منابع ۱۶۵

ضمائم ۱۷۱

چکیده انگلیسی ۱۹۴

فهرست جداول

جدول ..... صفحه

جدول ۴-۱ توزیع مشخصات فردی بیماران مبتلا به صرع ۱۲۳
جدول ۴-۲ توزیع مشخصات بیماری بیماران مبتلا به صرع ۱۲۳
جدول ۴-۳ توزیع مشخصات درمانی بیماران مبتلا به صرع ۱۲۴
جدول ۴-۴ توزیع فراوانی تعداد IED مشاهده شده در بیماران مبتلا به صرع ۱۲۴
جدول ۴-۵ مشخصات زمان کل برای IED مشاهده شده (بر حسب دقیقه) ۱۲۵
جدول ۴-۶ مشخصات فاصله زمانی بین IED مشاهده شده(بر حسب دقیقه) ۱۲۶
جدول ۴-۷ تعداد و میانگین فاصله زمانی بین IED مشاهده شده براساس گروه سنی ۱۲۹
جدول ۴-۸ تعداد و میانگین فاصله زمانی بین IED مشاهده شده براساس وضعیت جانبازی بیمار ۱۲۹
جدول ۴-۹ مدل های AG با برآورد واریانس به روش ساندویچی، PWP total time و PWP gap time ۱۳۱
جدول ۴-۱۰ مقادیر اولیه استفاده شده در مدل شکنندگی نرمال با دو بازه ۱۳۳
جدول ۴-۱۱ تعداد تکرارها و نقاط کوادراتور در مدل‌های شکنندگی نرمال با دو بازه ۱۳۳
جدول ۴-۱۲ نتیجه برازش مدل شکنندگی نرمال با دو بازه ۱۳۴
جدول ۴-۱۳ مقادیر اولیه استفاده شده در مدل شکنندگی برای فاصله زمانی با چهار بازه ۱۳۵
جدول ۴-۱۴ تعداد تکرارها و نقاط کوادراتور در مدل‌های شکنندگی برای فاصله زمانی با چهار بازه ۱۳۵
جدول ۴-۱۵ نتیجه برازش مدل شکنندگی نرمال با چهار بازه ۱۳۶
جدول ۴-۱۶ مقادیر اولیه استفاده شده در مدل اول شکنندگی وابسته به زمان با دو بازه ۱۳۷
جدول ۴-۱۷ تعداد تکرارها و نقاط کوادراتور در مدل اول شکنندگی وابسته به زمان با دو بازه ۱۳۷
جدول ۴-۱۸ نتیجه برازش مدل اول شکنندگی وابسته به زمان با دو بازه ۱۳۸
جدول ۴-۱۹ برآورد امیدریاضی و واریانس شکنندگی در مدل اول شکنندگی وابسته به زمان ۱۳۹
جدول ۴-۲۰ مقادیر اولیه استفاده شده در مدل اول شکنندگی وابسته به زمان با چهار بازه ۱۳۹
جدول ۴-۲۱ تعداد تکرارها و نقاط کوادراتور در مدل اول شکنندگی وابسته به زمان با چهار بازه ۱۴۰
جدول ۴-۲۲ نتیجه برازش مدل اول شکنندگی وابسته به زمان با چهار بازه ۱۴۰
جدول ۴-۲۳ برآورد واریانس شکنندگی در مدل اول شکنندگی وابسته به زمان ۱۴۱
جدول ۴-۲۴ مقادیر اولیه استفاده شده در مدل دوم شکنندگی وابسته به زمان با دو بازه ۱۴۲
جدول ۴-۲۵ تعداد تکرارها و نقاط کوادراتور در مدل دوم شکنندگی وابسته به زمان با دو بازه ۱۴۲
جدول ۴-۲۶ نتیجه برازش مدل دوم شکنندگی وابسته به زمان با دو بازه ۱۴۳
جدول ۴-۲۷ برآورد واریانس شکنندگی در مدل دوم شکنندگی وابسته به زمان ۱۴۳
جدول ۴-۲۸ مقادیر اولیه استفاده شده در مدل دوم شکنندگی وابسته به زمان با چهار بازه ۱۴۴
جدول ۴-۲۹ تعداد تکرارها و نقاط کوادراتور در مدل دوم شکنندگی وابسته به زمان با چهار بازه ۱۴۴
جدول ۴-۳۰ نتیجه برازش مدل دوم شکنندگی وابسته به زمان با چهار بازه ۱۴۵
جدول ۴-۳۱ برآورد واریانس شکنندگی در مدل دوم وابسته به زمان ۱۴۶

جدول ۴-۳۲ مقایسه مدل‌های حوادث بازگشتی در داده های صرع ۱۴۷
جدول ۴-۳۳ نتیجه شبیه سازی برای دو مدل شکنندگی وابسته به زمان ۱۵۳

فهرست شکل ها

شکل ..... صفحه

شکل ۱-۱ مثالی از سه فرد با حادثه بازگشتی ۴
شکل ۲-۱ نمایش بازه های خطر فاصله زمانی، زمان کل فرآیندشمارشی با داده های فرضی ۲۰
شکل ۲-۲ نمایش حادثه مشاهده شده و حادثه مشاهده نشده و زمان سانسور شدن برای هر فاصله زمانی . ۲۳
شکل ۲-۳ نمایش فاصله زمانی کامل و ناقص ۶۰
شکل ۳-۱ فلوچارت روش بهینه سازی نامقید ۸۵
شکل ۴-۱ فرآیند ثبت اطلاعات ۱۲۲
نمودار ۴-۱ نمودار بقای فاصله زمانی بین حوادث ۱۲۷

فصل اول

مقدمه و کلیات

۱–۱) مقدمه

در علوم کاربردی، داده‌های همبسته[1] یک عبارت کلی است که برای ساختار داده‌هایی مثل مشاهدات چند متغیره، داده‌های خوشه‌ای و اندازه‌های تکراری به کار می‌رود [۱]. وجه مشترک این نوع داده‌ها در این است که نوعی همبستگی در آنها مشاهده می‌شود. در مطالعات اپیدمیولوژی و بالینی محققان اغلب با داده‌های همبسته‌ای مواجه می‌شوند که پاسخ مورد بررسی در آنها از نوع بقا (زمان تا رخداد حادثه موردنظر) است لذا مساله آنها آنالیز چند متغیره داده‌های بقای همبسته است. همبستگی در داده‌های بقای همبسته به خاطر هم گروه شدن افراد در گروه‌ها (خوشه) یا تجربه چندین حادثه از یک نوع و یا حوادث کاملاً متفاوت توسط فرد در طول زمان باشد که هر یک آنالیزهای منحصر به خود را در پی دارند [۲].

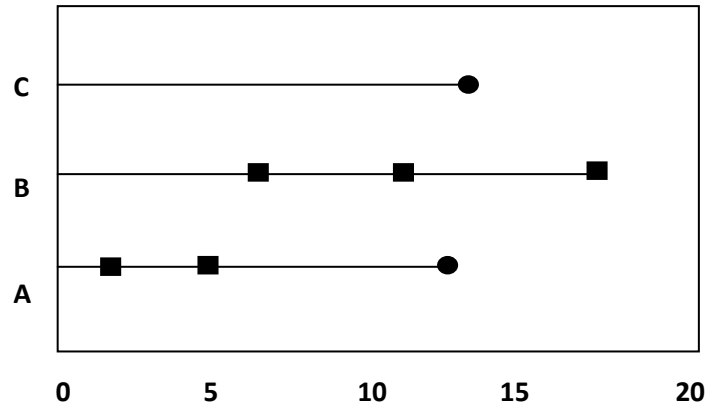
داده‌های حوادث خوشه ای[2] وقتی مشاهده می‌شوند که هر فرد یک حادثه را تجربه می‌کند و افراد به دلیل قرار گرفتن در یک گروه یا خوشه (مثل خانواده یا مدرسه) زمان وقوع حوادث به هم وابسته دارند [۳]. این وابستگی می‌تواند به دلایل شرایط محیطی و ژنتیکی مشترکی باشد که افراد آن خوشه آنها را تجربه می‌کنند [۴] و این شرایط از خوشه‌ای به خوشه دیگر متفاوت است. داده‌های حوادث چندگانه[3] وقتی مشاهده می‌شوند که هر فرد بیش از یک حادثه را تجربه کند لذا زمان این حوادث ذاتا به هم وابسته

[1] correlated data
[2] cluster events
[3] multivariate events

می‌شوند. داده‌های حوادث چندگانه خود می‌تواند به دو دسته حوادث چندگانه سری[1] و حوادث چندگانه موازی[2] که در آن حوادث از چندین نوع مثل عود تومور در قسمت های مختلف بدن (حوادث چند متغیره موازی نیز نامیده می شوند)، تقسیم شود [۳]. در داده‌های زمان شکست چندگانه سری، یک نوع حادثه چندین بار توسط بعضی یا همه افراد تحت مطالعه، مشاهده می‌شود [۳]. به این نوع داده، داده‌های حوادث بازگشتی[3] نیز می گویند (شکل۱-۱).

در شرایط پزشکی متعدد، افراد می‌توانند بروزی از بیماری را به صورت تکراری تجربه کنند مثل حملات آسم، عود تومور، حملات صرع، میگرن، عفونت، حملات قلبی، جراحت‌ها و بستری شدن در بیمارستان. فرآیند حوادث بازگشتی بخش مهم و وسیعی از کارها و تحقیقات پزشکی را شامل می‌شود [۶،۳]. جنبه مشترک در این نوع حوادث بازگشتی این است که حوادث دارای ترتیب ذاتی بوده و در طول زمان به صورت یک دنباله مشخص و پشت سر هم رخ می‌دهند. مشاهده حوادث بازگشتی می‌تواند قبل از اتمام مطالعه، پایان یابد. به عنوان مثال مشاهده بستری شدن گروهی از بیماران به دفعات پشت سرهم در بیمارستان با تمام شدن مدت مطالعه یا با از دست دادن بیمار در طول مطالعه یا مرگ بیمار پایان می‌یابد [۷،۶]. همبستگی در داده‌های حوادث بازگشتی به دو دلیل رخ می‌دهد: الف) ناهمگنی بین افراد[4]: در برخی مطالعات، بعضی افراد به خاطر عوامل غیرقابل اندازه‌گیری یا اندازه‌گیری نشده یا نامعلوم، حادثه‌ای را کمتر و یا بیشتر از دیگران تجربه می‌کنند، که این امر نوعی همبستگی درون-فردی را در تکرارها و در زمان وقوع حوادث آن فرد ایجاد می‌کند. ب) وابستگی حوادث (همبستگی درون-فردی)[5]: رخداد یک حادثه ممکن است احتمال حوادث بعدی را کمتر یا بیشتر کند. این وابستگی ممکن است به خاطر برخی مسائل بیولوژیکی باشد [۸] و باعث می‌شود که زمان بین حوادث برای بعضی از افراد کوتاه‌تر یا طولانی‌تر از بقیه افراد باشد یا به عبارتی دیگر در برخی حوادث زودتر و در برخی دیرتر رخ دهند [۹].

[1] serial event
[2] parallel event
[3] recurrent events
[4] heterogeneity across individuals
[5] within-subject correlation



شکل ۱-۱: مثالی از سه فرد با حادثه بازگشتی. علامت مربع توپر نشان‌دهنده وقوع حادثه و دایره توپر نشان‌دهنده سانسور است. فرد A دو حادثه دارد قبل از سانسورشدن. فرد B سه حادثه دارد که انتهای دوره یک حادثه رخ داده است. فرد C حادثه ای را قبل از سانسورشدن تجربه نکرده است.

طراحی مطالعه برای مطالعه طولی با ثبت حوادث بازگشتی به دو صورت می‌تواند باشد. طراحی اول اینگونه است که رخداد اصلی حادثه موردنظر معیار ورود است و سپس رخدادهای بعدی همان حادثه در طول مدت پیگیری مشاهده می‌شوند. در این حالت زمان رخداد حادثه اول به عنوان زمان مبدا (صفر) در نظر گرفته می‌شود. در طراحی دوم افراد از جامعه هدف نمونه گیری شده و حوادث در طول دوره پیگیری مشاهده می‌گردند. مبدا این نوع طرح می‌تواند زمان ورود و ثبت نام در مطالعه یا زمان شروع مداخله باشد. در این حالت رخداد حادثه اصلی برای ورود نمونه لازم نیست و در نتیجه ممکن است که هیچ حادثه‌ای در طول دوره مشاهده نشود [۱۰].

یکی از موضوعاتی که در داده‌های حوادث بازگشتی مطرح می‌شود این است که چه مقیاسی از زمان وارد مدل شود. مقیاس‌های زمانی که معمولا در آنالیز استفاده می‌شود شامل زمان کل[1] و فاصله زمانی بین حوادث[2] است. منظور از زمان کل زمان بین آغاز مطالعه برای فرد تا زمان وقوع حادثه مورد نظر است و فاصله زمانی بین حوادث برابر با زمان از حادثه بلافاصله قبلی تا حادثه بعدی است. در نظر گرفتن فواصل زمانی بین حوادث در مدل سازی اخیرا بیشتر مورد توجه قرار گرفته و جذاب‌تر است چرا که به بعضی سوالات جالب می‌تواند پاسخ بدهد مثل توزیع زمان تا عود بعدی بیماری در بیمارانی که قبلاً عود دیگری را تجربه کرده‌اند [۱۱].

آنالیز داده‌های حوادث بازگشتی (عودکننده) مرتباً در مطالعات اپیدمیولوژی و بالینی نیاز می‌شود و هدف اکثر این آنالیزها ارزیابی اثر درمان بر به تاخیر انداختن عود بیماری و طولانی کردن بقای بیمار است. به دلیل وجود این همبستگی ساختار داده‌ها پیچیده شده لذا نیاز به روش‌های آماری منطقی و مناسب برای پاسخ به اینگونه سوالات تحقیق دارد [۱۲].

۱–۲) بیان مساله و ضرورت اجرا

داده‌های حادثه بازگشتی به خاطر شکل طولی بودن و وابستگی بین حوادث، جنبه‌های مخصوص به خود را داشته و ساختار آماری متفاوتی از زمان‌های بقای چند متغیره موازی دارند. در اینگونه داده‌ها اغلب زمان وقوع بیماری (یا علائم بالینی خاص) همراه با کووریت های شخصی و محیطی ثبت می‌شود و هدف آنها بررسی اثر این عوامل بر زمان وقوع حوادث است. در بین آنالیزهای بقای موجود، مدل خطرات متناسب کاکس یک ابزار مفید در بررسی عوامل موثر بر بقا بوده و بیشتر از سایر مدل‌ها استفاده می‌شود ولی این مدل در مطالعاتی به کار می‌رود که حادثه تنها یک بار رخ می‌دهد و پیش فرض اصلی آن استقلال داده‌ها است. لذا استفاده از این روش برای داده‌های وابسته، منجر به برآورد واریانس ناصحیح و

[1] total time
[2] gap time

درنتیجه استنباط‌های غلط می‌گردد لذا برای آنالیز این نوع داده‌ها بایستی روش‌های دیگری استفاده شود. [۱۳،۱۲]

محققان اغلب از روش‌های ساده‌تر برای آنالیز زمان وقوع اتفاقات مربوط به سلامت که به صورت تکراری رخ می‌دهد استفاده می‌کنند. آنالیزهای معمولی که بر روی چنین داده‌هایی اغلب انجام می‌شود بر اساس فراوانی حوادث، زمان تا حادثه اول، زمان بقای کلی یا مدل‌های جداگانه برای هر حادثه که به صورت مستقل از هم درنظر گرفته می‌شوند، است که روش‌های ناکافی هستند. به عنوان مثال بررسی توزیع زمان بقای بدون بیماری[1] یعنی زمان از ابتدای شروع بیماری تا اولین حادثه که به علت نادیده گرفتن بقیه عودهای بیماری یا حتی مرگ ناشی از بیماری که خود یک عود بیماری به شمار می‌رود، باعث خدشه دار شدن و کم شدن توان آماری نتایج می‌شود [۱۶،۴]. همچنین ممکن است اثر متغیرهای مستقل بر تعداد حوادث به وقوع پیوسته برای هر فرد از طریق مدل‌های خطی تعمیم یافته با توزیع پواسون بررسی گردد ولی در این روش زمان وقوع حوادث و روابط بین آنها لحاظ نمی شود [۱۰].

برای داده‌های حوادث تکراری در طول یک دوره مشخص، روش‌های آنالیزی پیچیده‌تری لازم است تا برآوردهای دقیق‌تر و استنباط‌های کاراتری بدست آید. یک بحث مهم در آنالیز این نوع داده‌ها نحوه لحاظ نمودن وابستگی حوادث رخ داده برای یک نمونه در مدل و تعدیل همبستگی می‌باشد [۹]. نتیجه در نظر نگرفتن وابستگی این حوادث یعنی استفاده از روش‌های درستنمایی جزئی مرسوم (با پیش فرض استقلال) برای داده‌های وابسته، برآوردهای نااریبی برای ضرایب رگرسیونی فراهم می‌آورد اما برآورد واریانس ها درست نیست. محققان از طریق شبیه سازی نشان داده‌اند که سطح وابستگی بین حوادث اثر معنی‌داری بر روی اریبی برآورد ضرایب رگرسیونی ندارد ولی برآورد واریانس صحیح نیست [۱۳-۱۵]. تاکنون مدل‌های مختلفی برای برازش حوادث بازگشتی پیشنهاد شده است که به صورت تعمیمی از مدل کاکس هستند از جمله مدل‌های حاشیه‌ای، مدل‌های تصحیح واریانس[2]، مدل‌های شرطی و مدل‌های

[1] disease-free survival
[2] variance-corrected model

شکنندگی که در آنها همبستگی وارد مدل می‌گردد [۸]. اخیرا این مدل‌ها برای زمان‌های بقای همبسته کاربرد بیشتری یافته است. در بعضی از این مدل‌ها، هدف بدست آوردن برآورد پارامترهای رگرسیونی است در حالیکه در بعضی موارد هدف برآورد همبستگی است [۱۶]. شبیه سازی نشان داده است که در روش درستنمایی جزئی استاندارد با افزایش وابستگی بین داده‌ها، خطای نوع اول نیز افزایش می‌یابد ولی در روش‌های حاشیه‌ای و شرطی که برای داده‌های وابسته طراحی شده اند، در تمام وابستگی ها، مقدار خطای نوع اول قابل قبول است [۱۳].

البته هر یک از این مدل‌ها نیز دارای محدودیت‌ها و مشکلاتی هستند. به عنوان مثال اگر مداخله در پیشگیری حادثه اول موثر باشد ولی بر روی حوادث بعدی اثر نداشته باشد، مدل حاشیه‌ای نشان می‌دهد که این مداخله بر روی همه حوادث موثر است. علت نیز این می‌باشد که اگر درمان، حادثه اول را به تاخیر بیاندازد در نتیجه حوادث بعدی نیز به خودی خود عقب میافتد و مدت آنها بیشتر می‌شود حتی اگر درمان موثر نباشد. در مدل‌های تصحیح واریانس نیز محققانی از جمله کلی[۱] و لیم[۲] (۲۰۰۰) با شبیه سازی نشان دادند که استفاده از واریانس استوار[۳] برای لحاظ کردن همبستگی درون فردی کافی نیست در این حالت پیشنهاد می‌شود که به جای استفاده از مدل‌هایی که تصحیح واریانس در آنها انجام می‌شود از اثرات تصادفی (شکنندگی) در مدل استفاده شود، گرچه مدل پیچیده‌تر می‌گردد [۳].

در مدل‌های شکنندگی اثرات فردی (شکنندگی) وارد مدل می‌شود. منظور از اثرات فردی این است که افراد عوامل خطر متفاوتی نسبت به هم دارند که البته این اثرات فردی با کووریت‌های مشاهده شده شرح داده نمی‌شوند [۵]. نادیده گرفتن شکنندگی در صورتی که وجود داشته باشد منجر به کم برآورده[۴] شدن اثر کووریت ها و دقت کم در مدل بقا می‌شود. در نتیجه در روش‌های حاشیه‌ای باید توجه بیشتری صورت گیرد تا مدل معتبر باشد [۶]. در مدل‌های شکنندگی فرض می‌شود که زمان حوادث به شرط

[1] Kelly
[2] Lim
[3] robust
[4] underestimate

جملات شکنندگی مشاهده نشده مستقل هستند. مدل های شکنندگی برای بیان اثرات فردی بهتر است گرچه می‌تواند ناکافی باشد؛ چرا که در اکثر موارد فرض می‌شود اثرات فردی تصادفی (شکنندگی) در طول مدت پیگیری ثابت است ولی در کارهای عملی این پیش فرض محدودیت‌هایی دارد. تجربه نشان می‌دهد که اغلب هر چه فاصله بین حوادث بیشتر شده و حوادث دورتر از هم می شوند، وابستگی آنها نیز کمتر می شود همچنین شکنندگی یا اثرات فردی در طول زمان کاهش می‌یابد و یا بعد از شکست اول افزایش می یابد لذا مدل‌هایی با شکنندگی های وابسته به زمان، مدل‌های واقعی‌تری هستند [۱۵].

در این مطالعه فرض می‌شود که شکنندگی ثابت نیست و با زمان تغییر می‌کند. از طرفی با توجه به اینکه فاصله زمانی بین حوادث از نظر بالینی مهم بوده و در تصمیم گیریها بسیار موردتوجه قرار می‌گیرد، چرا که فاصله زمانی بین حوادث می‌تواند نشان دهنده شدت بیماری یا ابتلا به بیماری خاص یا تعیین مدت زمان لازم جهت پیگیری بیماری باشد، در این طرح اثر متغیرهای مستقل و اثر تصادفی وابسته به زمان بر روی فواصل زمانی بین حوادث بررسی می‌گردد. مبحث دیگری که در این طرح به آن توجه می-شود روش برآورد پارامترها در مدل شکنندگی وابسته به زمان است. روش‌هایی که اکنون برای برآورد پارامترهای مدل شکنندگی استفاده می شود (روش برآورد EM، روش برآورد ناپارامتری، روش بیزی و ...) روش‌های سخت و پیچیده و دارای محدودیت و زمان‌بر هستند که باعث محدود شدن کاربرد این نوع مدل‌ها در عمل می‌شود. به عنوان مثال در روش الگوریتم EM به این دلیل که اکثر امیدهای شرطی اثرات تصادفی به شرط داده‌های مشاهده شده شکل بسته ای ندارند و بایستی روش مونت کارلو در گام E برای تقریب این جملات استفاده شود لذا از نظر محاسباتی خیلی سنگین و برنامه نویسی آن سخت می‌شود. همچنین همگرایی الگوریتم به آهستگی انجام می‌شود و برای بدست آوردن خطای استاندارد برآوردها باید عملیات زیادی را انجام داد. روش درستنمایی ناپارامتری PPL[1] سریع بوده و در اکثر برنامه های آماری مثل S-plus و R وجود دارد ولی خطای استاندارد شکنندگی در آنها به طور مستقیم قابل برآورد نیست و

[1] penalized partial likelihood