

الله
البر
الرحمن
الرحيم



تاییدیه اعضای هیات داوران حاضر در جلسه دفاع از پایان نامه کارشناسی ارشد

آقای سامان معروفی زاده دانشجوی رشته آمارزیستی پایان نامه کارشناسی ارشد خود را با عنوان: به کارگیری مدل‌های رگرسیونی خطرات جمعی در مطالعه زمان بقای بیماران مبتلا به سرطان معده در تاریخ ۱۳۹۰/۴/۱ ارائه کردند.

بدینوسیله اعضای هیات داوران نسخه نهایی این پایان نامه را از نظر فرم و محتوا تایید کرده و پذیرش آنرا برای تکمیل درجه کارشناسی ارشد پیشنهاد می‌کنند.

نام و نام خانوادگی و امضاء اعضای هیأت داوران:

دکتر ابراهیم حاجی زاده (استاد راهنما)

دکتر احمدرضا باغستانی (استاد مشاور)

دکتر اکبر بیگلریان (استاد ناظر)

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

آئین 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) ی 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:

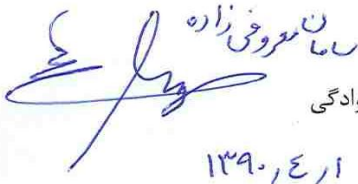
" کتاب حاضر، حاصل پایان نامه کارشناسی ارشد نگارنده در رشته آمار زیستی است که در سال ۹۰ در دانشکده علوم پزشکی دانشگاه تربیت مدرس به راهنمایی دکتر ابراهیم حاجی زاده، مشاوره دکتر احمد رضا باغستانی از آن دفاع شده است.

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 سامان معروفی زاده 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
امضا
۱۳۹۰، ۴، ۱

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 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 (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 آیین‌نامه‌های 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 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 سامان معروفی زاده دانشجوی رشته آمار زیستی ورودی سال تحصیلی ۸۷ مقطع کارشناسی ارشد دانشکده علوم پزشکی متعهد می شوم کلیه نکات مندرج در آیین‌نامه 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.»

سامان معروفی زاده
امضا
تاریخ

۱۳۹۰، ۴، ۱



دانشگاه تربیت مدرس
دانشکده علوم پزشکی

پایان نامه

دوره کارشناسی ارشد در رشته آمار زیستی

عنوان

به کارگیری مدل های رگرسیونی خطرات جمعی در مطالعه زمان بقای بیماران
مبتلا به سرطان معده

نگارش

سامان معروفی زاده

استاد راهنما

دکتر ابراهیم حاجی زاده

استاد مشاور

دکتر احمد رضا باغستانی

تابستان ۱۳۹۰

تقدیم به

مادر مهربانم

و

پدر بزرگوارم

تشکر و قدردانی

اکنون که به این مرحله از فراگیری دانش رسیدم، از تمام اساتید و بزرگوارانی که در هدایت و انجام مراحل تحقیق اینجانب را یاری نموده‌اند تشکر و قدردانی می‌کنم.

از استاد بزرگوار و ارجمند جناب آقای دکتر ابراهیم حاجی‌زاده به خاطر راهنمایی و هدایت پایان‌نامه.

از استاد مشاور عزیز جناب آقای دکتر احمدرضا باغستانی که در امر مشاوره پایان‌نامه کمک و یاری نمودند.

از اساتید محترم و بزرگوار گروه آمار زیستی جناب آقایان دکتر سقراط فقیه‌زاده و دکتر انوشیروان کاظم‌نژاد که در طول تحصیل از محضرشان بهرمنند شدم.

از هم‌کلاسی‌های دوره تحصیل سرکار خانم‌ها محدثه شجاعی و آزاده قاهری و جناب آقایان شهرام ارسنگ و روح‌الله روهنده که از مصاحبت و همفکری ایشان بهره‌مند بوده‌ام.

چکیده

مدل مخاطرات متناسب کاکس، یک مدل ضربی در تحلیل بقا است و اغلب در مطالعات زیست پزشکی مورد استفاده قرار می‌گیرد. یک پیش فرض مهم و اساسی در این مدل، پذیره متناسب بودن خطر برای تمام متغیرهای مستقل موجود در مدل نهایی می‌باشد. وقتی پیش فرض خطرات متناسب سؤال برانگیز باشد، یک مدل جایگزین، مدل‌های رگرسیونی مخاطرات جمعی است. در مدل‌های رگرسیونی مخاطرات جمعی از قبیل مدل آلن و مدل لین-یانگ، متغیرهای کمکی به صورت جمعی بر تابع خطر پایه عمل می‌کند. در این مدل‌ها ارتباط بین متغیرهای کمکی و متغیر پاسخ بر حسب تفاوت خطر یا خطر اضافی بیان می‌شود در حالی که در مدل کاکس بر حسب نسبت خطر بیان می‌شود. مدل مخاطرات جمعی آلن می‌تواند اطلاعات دقیق و کامل‌تری در خصوص اثر یک عامل خطر در طول زمان ارائه کند.

هدف مطالعه حاضر، به کارگیری مدل‌های رگرسیونی مخاطرات جمعی در تحلیل بقای بیماران مبتلا به سرطان معده و مقایسه نتایج به دست آمده با مدل کاکس می‌باشد. در این مطالعه همگروهی تاریخی به ۲۱۳ پرونده بیمار مبتلا به سرطان معده که از بهمن ۱۳۸۲ لغایت دی‌ماه ۱۳۸۷ در بخش گوارش بیمارستان طالقانی تهران تحت درمان بودند؛ مراجعه شد و با تماس تلفنی اطلاعات مربوط به بقای بیماران جمع‌آوری گردید. وضعیت بقای بیماران بر حسب مدل کاکس و مدل‌های رگرسیونی مخاطرات جمعی با در نظر گرفتن متغیرهایی مانند سن هنگام تشخیص، جنسیت، وجود متاستاز، اندازه تومور، نوع هیستوپاتولوژی، متاستاز غدد لنفاوی و مرحله پاتولوژی تحلیل شد. همچنین نمودار ضرایب رگرسیونی تجمعی (نمودار آلن) جهت بررسی اثر متغیرهای مستقل در طول زمان رسم شد. میزان بقای پنج ساله بیماران مورد بررسی ۱۴/۶ درصد و میانه طول عمر در این بررسی ۲۹/۶ ماه به دست آمد. نسبت مرد به زن در بین ۲۱۳ بیمار ۱: ۲/۶۱ بود. تحلیل‌های چندمتغیری بر حسب مدل‌های جمعی و مدل کاکس نشان داد که سن هنگام تشخیص، اندازه تومور و مرحله پاتولوژی با طول عمر بیماران ارتباط معنی‌دار داشتند ($p < 0.05$). به علاوه، بر طبق نمودار آلن مرحله پاتولوژی دارای اثر تأخیری بود. عوامل دیگر اثر معنی‌داری بر زمان بقاء نداشتند ($p > 0.05$).

علی‌رغم تمایل بسیاری از محققین در استفاده از مدل کاکس در تحلیل داده‌های بقا، مدل‌های رگرسیونی مخاطرات جمعی دید جدیدی به محقق در مطالعات زیست‌پزشکی می‌دهد؛ به علاوه مدل آلن این قابلیت را دارد که اثر متغیرهای مستقل را در طول زمان بررسی کند. این مطالعه نشان داد که تشخیص سرطان معده در سنین پایین‌تر و مراحل اولیه بیماری، منجر به افزایش قابل توجهی در میزان بقای بیماران می‌گردد.

واژگان کلیدی: سرطان معده، مدل‌های مخاطرات جمعی، مدل کاکس، بقا، عامل پیش‌آگهی کننده

فهرست مطالب

۱	فصل اول: مقدمه و مروری بر مطالعات گذشته
۲	۱-۱. مقدمه
۳	۲-۱. بیان مسئله
۵	۳-۱. مدل‌های رگرسیونی مخاطرات جمعی
۶	۴-۱. مروری بر مطالعات انجام شده در زمینه مدل‌های رگرسیونی مخاطرات جمعی
۸	۵-۱. سرطان
۹	۶-۱. نحوه ایجاد سرطان
۱۰	۷-۱. اپیدمیولوژی سرطان
۱۰	۱-۷-۱. اپیدمیولوژی سرطان در جهان
۱۱	۲-۷-۱. اپیدمیولوژی سرطان در ایران
۱۱	۸-۱. سرطان معده
۱۲	۱-۸-۱. انواع سرطان معده
۱۲	۲-۸-۱. علائم سرطان معده
۱۳	۳-۸-۱. عوامل خطر سرطان معده
۱۴	۴-۸-۱. مراحل سرطان معده
۱۴	۵-۸-۱. تشخیص سرطان معده
۱۴	۶-۸-۱. درمان سرطان معده
۱۵	۷-۸-۱. پیشگیری
۱۶	۹-۱. اپیدمیولوژی سرطان معده
۱۶	۱-۹-۱. اپیدمیولوژی سرطان معده در جهان
۱۶	۲-۹-۱. اپیدمیولوژی سرطان معده در ایران
۱۷	۱۰-۱. اهداف پژوهش
۱۷	۱-۱۰-۱. هدف اصلی
۱۷	۲-۱۰-۱. اهداف اختصاصی
۱۷	۱۱-۱. فرضیه‌ها
۱۸	فصل دوم: مواد و روش‌ها
۱۹	۱-۲. مقدمه
۱۹	۲-۲. روش‌شناسی پژوهش
۱۹	۱-۲-۲. جامعه، نمونه و نوع مطالعه

۲۰ ۲-۲-۲. متغیرهای مورد بررسی
۲۰ ۳-۲-۲. تجزیه و تحلیل داده‌ها
۲۱ ۳-۲-۳. مقدمات تحلیل بقا
۲۱ ۱-۳-۲. شاخص‌های تحلیل بقا
۲۱ ۱-۱-۳-۲. تابع بقا
۲۲ ۲-۱-۳-۲. تابع خطر
۲۲ ۲-۳-۲. سانسور شدن
۲۳ ۱-۲-۳-۲. سانسور شده از راست
۲۳ ۲-۲-۳-۲. سانسور شده از چپ
۲۳ ۳-۲-۳-۲. سانسور شده فاصله‌ای
۲۳ ۴-۲. روش‌های متداول در تحلیل داده‌های بقا
۲۴ ۱-۴-۲. روش کاپلان-مایر
۲۴ ۲-۴-۲. آزمون‌های ناپارامتری برای مقایسه منحنی‌های بقا
۲۵ ۱-۲-۴-۲. آزمون‌های ناپارامتری برای مقایسه منحنی‌های بقا در دو گروه
۲۷ ۲-۲-۴-۲. آزمون‌های ناپارامتری برای مقایسه منحنی‌های بقا در چند گروه
۲۹ ۳-۴-۲. مدل مخاطرات متناسب کاکس
۲۹ ۱-۳-۴-۲. توصیف مدل مخاطرات متناسب کاکس
۳۱ ۲-۳-۴-۲. آزمون پارامترهای مدل مخاطرات متناسب کاکس
۳۲ ۳-۳-۴-۲. بررسی پیش‌فرض مخاطرات متناسب و برازش کلی مدل کاکس
۳۳ ۵-۲. مدل‌های رگرسیونی مخاطرات جمعی
۳۳ ۱-۵-۲. مدل مخاطرات جمعی آلن
۳۳ ۱-۱-۵-۲. توصیف مدل مخاطرات جمعی آلن
۳۶ ۲-۱-۵-۲. آزمون فرضیه
۳۷ ۳-۱-۵-۲. ارزیابی برازش مدل مخاطرات جمعی
۴۰ ۴-۱-۵-۲. ویژگی‌های مدل آلن
۴۱ ۲-۵-۲. مدل مخاطرات جمعی لین-یانگ
۴۴ فصل سوم: نتایج و یافته‌ها
۴۵ ۱-۳. مقدمه
۴۵ ۲-۳. خصوصیات بالینی و پاتولوژیک بیماران مبتلا به سرطان معده
۴۷ ۳-۳. تحلیل عوامل مرتبط با بقای بیماران با استفاده از روش کاپلان-مایر و آزمون لگ-رتبه
۵۲ ۴-۳. تحلیل عوامل مرتبط با بقای بیماران با استفاده از مدل مخاطرات متناسب کاکس

۵۵ ۵-۳ بررسی پیش فرض خطرات متناسب و برازش کلی مدل کاکس
۵۷ ۶-۳ تحلیل عوامل مرتبط با بقای بیماران با استفاده از مدل مخاطرات جمعی آلن
۵۸ ۷-۳ نمودارهای ضرایب رگرسیون تجمعی (نمودارهای آلن)
۵۸ ۸-۳ ارزیابی برازش مدل مخاطره جمعی
۵۸ ۹-۳ تحلیل عوامل مرتبط با بقای بیماران با استفاده از مدل مخاطرات جمعی لین-یانگ

۶۸ فصل چهارم: بحث، نتیجه گیری و پیشنهادها
۶۹ ۱-۴ مقدمه
۷۰ ۲-۴ بحث
۷۳ ۳-۴ نتیجه گیری
۷۴ ۴-۴ پیشنهادها
۷۶ فهرست منابع
۸۳ چکیده انگلیسی

فهرست جداول

- جدول ۱-۱. تقسیم‌بندی انواع سرطان معده ۱۲
- جدول ۱-۳. خصوصیات بالینی و پاتولوژیک بیماران مبتلا به سرطان معده و نتایج آزمون کای دو ۴۶
- جدول ۲-۳. نتایج آزمون لگ-رتبه جهت مقایسه منحنی‌های بقا در سطوح مختلف عوامل بالینی و پاتولوژیک ... ۴۸
- جدول ۳-۳. نتایج حاصل از تحلیل تک‌متغیری عوامل مرتبط با بقای بیماران مبتلا به سرطان معده با استفاده از مدل مخاطرات متناسب کاکس ۵۳
- جدول ۴-۳. نتایج حاصل از تحلیل چندمتغیری عوامل مرتبط با بقای بیماران مبتلا به سرطان معده با استفاده از مدل مخاطرات متناسب کاکس ۵۴
- جدول ۵-۳. بررسی پیش فرض خطرات متناسب با استفاده از باقیمانده‌های اسکنفیلد ۵۵
- جدول ۶-۳. نتایج حاصل از تحلیل تک‌متغیری عوامل مرتبط با بقای بیماران مبتلا به سرطان معده با استفاده از مدل مخاطرات جمعی آلن ۵۷
- جدول ۷-۳. نتایج حاصل از تحلیل چندمتغیری عوامل مرتبط با بقای بیماران مبتلا به سرطان معده با استفاده از مدل مخاطرات جمعی آلن ۵۸
- جدول ۸-۳. نتایج حاصل از تحلیل تک‌متغیری عوامل مرتبط با بقای بیماران مبتلا به سرطان معده با استفاده از مدل مخاطرات جمعی لین-یانگ ۶۶
- جدول ۹-۳. نتایج حاصل از تحلیل چندمتغیری عوامل مرتبط با بقای بیماران مبتلا به سرطان معده با استفاده از مدل مخاطرات جمعی لین-یانگ ۶۷

فهرست نمودارها

- نمودار ۳-۱. تابع بقا بیماران مبتلا به سرطان معده ۴۷
- نمودار ۳-۲. منحنی‌های بقا بیماران مبتلا به سرطان معده بر حسب جنسیت ۴۹
- نمودار ۳-۳. منحنی‌های بقا بیماران مبتلا به سرطان معده بر حسب وضعیت متاستاز ۴۹
- نمودار ۳-۴. منحنی‌های بقا بیماران مبتلا به سرطان معده بر حسب اندازه تومور ۵۰
- نمودار ۳-۵. منحنی‌های بقا بیماران مبتلا به سرطان معده بر حسب نوع هیستوپاتولوژی ۵۰
- نمودار ۳-۶. منحنی‌های بقا بیماران مبتلا به سرطان معده بر حسب متاستاز غدد لنفاوی ۵۱
- نمودار ۳-۷. منحنی‌های بقا بیماران مبتلا به سرطان معده بر حسب مرحله پاتولوژی ۵۱
- نمودار ۳-۸. نمودار باقیمانده‌های کاکس-اسنل برای ارزیابی برازش کلی مدل مخاطرات متناسب کاکس (حالت چند متغیره) ۵۶
- نمودار ۳-۹. ارزیابی شکل تابعی سن هنگام تشخیص در بیماران مبتلا به سرطان معده با استفاده از نمودار باقیمانده-های مارتینگل ۵۶
- نمودار ۳-۱۰. برآورد میزان مخاطره تجمعی مبنا و فاصله اطمینان ۹۵ درصد برای بیماران مبتلا به سرطان معده .. ۵۹
- نمودار ۳-۱۱. برآورد اثر تجمعی سن هنگام تشخیص بیماری و فاصله اطمینان ۹۵ درصد برای بیماران مبتلا به سرطان معده ۶۰
- نمودار ۳-۱۲. برآورد خطر تجمعی اضافی مردان در مقایسه با زنان و فاصله اطمینان ۹۵ درصد در بیماران مبتلا به سرطان معده ۶۰
- نمودار ۳-۱۳. برآورد خطر تجمعی اضافی بیمارانی که دچار متاستاز بودند در مقایسه با بیمارانی که متاستاز نداشتند و فاصله اطمینان ۹۵ درصد در بیماران مبتلا به سرطان معده ۶۱
- نمودار ۳-۱۴. برآورد خطر تجمعی بیماران با اندازه تومور بزرگتر از ۳۵ میلی‌متر در مقایسه با بیماران با اندازه تومور کوچکتر از ۳۵ میلی‌متر و فاصله اطمینان ۹۵ درصد در بیماران مبتلا به سرطان معده ۶۱
- نمودار ۳-۱۵. برآورد خطر تجمعی اضافی بیماران با هیستوپاتولوژی سلول‌های نگین انگشتی در مقایسه با بیماران با هیستوپاتولوژی آدنوکارسینوما و فاصله اطمینان ۹۵ درصد در بیماران مبتلا به سرطان معده ۶۲
- نمودار ۳-۱۶. برآورد خطر تجمعی اضافی بیماران با هیستوپاتولوژی سایر در مقایسه با بیماران با هیستوپاتولوژی آدنوکارسینوما و فاصله اطمینان ۹۵ درصد در بیماران مبتلا به سرطان معده ۶۲
- نمودار ۳-۱۷. برآورد خطر تجمعی اضافی بیماران با متاستاز غدد لنفاوی N2 در مقایسه با بیماران با متاستاز غدد لنفاوی N1 و فاصله اطمینان ۹۵ درصد در بیماران مبتلا به سرطان معده ۶۳

نمودار ۳-۱۸. برآورد خطر تجمعی اضافی بیماران با متاستاز غدد لنفاوی N3 در مقایسه با بیماران با متاستاز غدد لنفاوی N2 و فاصله اطمینان ۹۵ درصد در بیماران مبتلا به سرطان معده ۶۳

نمودار ۳-۱۹. برآورد خطر تجمعی مرحله پاتولوژی پیشرفته در مقایسه با مرحله پاتولوژی اولیه و فاصله اطمینان ۹۵ درصد در بیماران مبتلا به سرطان معده ۶۴

نمودار ۳-۲۰. نمودار آرجس برای بررسی کفایت مدل مخاطره جمعی (سن کمتر از ۵۸/۶ سال —، سن بیشتر از ۵۸/۶ سال ---) ۶۵

نمودار ۳-۲۱. نمودار فرآیند باقیمانده مارتینگل و حدود اطمینان زوجی ۹۵ درصد برای گروه سنی کمتر از ۵۸/۶ سال ۶۵

فصل اول

مقدمه و

مروری بر مطالعات گذشته

۱-۱. مقدمه

بررسی طول عمر و احتمالات بقا امروزه قدمتی چنددهه‌ای یافته است. محاسبه میانه طول عمر در گروه‌های مختلف جامعه انسانی و بررسی عوامل مؤثر بر آن جهت دستیابی به راه‌هایی برای افزایش سلامت افراد و طول عمر آنان یکی از دغدغه‌های هر روزه دانشمندان شاخه‌های مختلف علوم زیستی و پزشکی است. به هر حال طول عمر افراد در هر جامعه آماری، متغیری تصادفی است و لذا پیش‌بینی آن جز از طریق استفاده از روش‌های آماری امکان‌پذیر نیست. به همین دلیل است که در سال‌های اخیر به ویژه از دهه هفتاد میلادی متخصصین علم آمار زیستی تلاش بسیاری در این زمینه داشته‌اند. داده‌های بقا^۱ دارای ویژگی‌های خاصی هستند که نمی‌توان از روش‌های متداول آماری جهت تجزیه و تحلیل آنها استفاده نمود، که از آن جمله می‌توان به سانسور شدن^۲ (ناتمام بودن) برخی از داده‌ها و بریدگی^۳ اشاره نمود [۱ و ۲].

یکی از محورهای اصلی پژوهش در عرصه سلامت و بیماری، بررسی طول عمر و یا زمان بازگشت بیماری در بیمارانی است که درمانی را دریافت کرده‌اند، و به این ترتیب است که ارزیابی اثربخشی درمان‌ها ممکن می‌شود. البته از این روش‌ها در سایر مطالعات پزشکی و زیستی نیز به وفور استفاده می‌شود.

از جمله روش‌های ارائه شده در این رابطه، تجزیه و تحلیل بقا با استفاده از مدل‌های رگرسیونی مخاطرات جمعی^۴ است که به بررسی میزان بقای بیماران و عوامل مؤثر بر این کمیت می‌پردازد.

¹ Survival data

² Censoring

³ Truncation

⁴ Additive hazards regression models

مطالعات اندکی در خصوص چگونگی استفاده از مدل‌های رگرسیونی مخاطرات جمعی در تحقیقات پزشکی به خصوص در زمینه تحلیل بقا که ویژگی اصلی آنها ناتمامی داده‌ها است، صورت گرفته است ولی در سال‌های اخیر استفاده از این مدل‌ها به خصوص مدل مخاطرات جمعی آلن با توجه به انعطاف‌پذیری کافی این مدل برای ورود متغیرهای مستقل وابسته به زمان و قابلیت بررسی اثر متغیرهای مستقل در طول زمان با استفاده از نمودار ضرایب رگرسیون تجمعی^۱ (نمودار آلن^۲) رواج یافته است.

۱-۲. بیان مسئله

تحلیل بقا مجموعه‌ای از روش‌های آماری برای تحلیل داده‌هایی است که متغیر پاسخ آنها زمان لازم تا رخداد یک پیشامد است. منظور از زمان در این تعریف می‌تواند سال، ماه، هفته و یا زمان شروع یک مطالعه تا زمان رخداد پیشامد مورد نظر باشد و یا سن فرد در زمان رخداد پیشامد مورد نظر باشد. منظور از پیشامد مورد نظر می‌تواند مرگ، بروز بیماری، عود بیماری، فروکش کردن نشانه‌های بیماری، بهبودی و یا هر تجربه تعریف شده‌ای باشد که افراد با آن مواجه می‌شوند. در تحلیل بقا معمولاً به متغیر زمان، عنوان زمان بقا^۳ داده می‌شود چون بقای هر یک از افراد تحت مطالعه بررسی می‌شود و نیز به پیشامد مورد نظر عنوان شکست^۴ داده می‌شود به دلیل اینکه نوع پیشامد مورد نظر معمولاً مرگ، بروز بیماری یا برخی دیگر از تجربه‌های منفی برای افراد است. البته ممکن است پیشامد مورد نظر یک تجربه مثبت نیز باشد؛ مثلاً بازگشت به کار بعد از انجام یک عمل جراحی. در این تحلیل هر مشاهده دارای یک نقطه شروع، یک نقطه پایانی و یک واحد زمانی تعریف شده برای فاصله بین این دو نقطه است. به طور کلی اهداف تحلیل بقا عبارتند از: برآورد و تفسیر توابع بقا یا خطر، مقایسه توابع بقا یا خطر و ارزیابی رابطه بین متغیرهای توضیحی و زمان بقا [۱ و ۲].

¹ Cumulative regression coefficients

² Aalean's plot

³ Survival Time

⁴ Failure

در مدل‌سازی داده‌های بقا، روش‌های مختلفی جهت تعیین عوامل مؤثر بر زمان بقا وجود دارد. از جمله این روش‌ها می‌توان به جدول عمر^۱، برآوردگر کاپلان-مایر^۲ و مدل‌های رگرسیونی بقا اشاره کرد. در بررسی داده‌های بقا اگر هدف، توصیف زمان بقا بدون در نظر گرفتن متغیرهای کمکی باشد از روش‌های ناپارامتری مانند جدول عمر و کاپلان-مایر استفاده می‌شود. ولی اگر هدف، بررسی اثرات متغیرهای کمکی بر روی زمان بقا باشد مدل‌های رگرسیونی بقا مورد استفاده قرار می‌گیرد [۱ و ۲]. به طور کلی دو نوع مدل رگرسیونی برای داده‌های بقا وجود دارد: مدل مخاطرات متناسب کاکس^۳ به عنوان یک روش نیمه‌پارامتری [۳] و مدل‌های شتاب‌دار زمان شکست^۴ از قبیل مدل وایبل^۵، نمایی^۶ و لگ نرمال^۷ و ... به عنوان روش‌های پارامتری [۱].

در تحلیل بقا مدل مخاطرات متناسب کاکس، یک مدل ضربی^۸ است و اغلب در مطالعات زیست‌پزشکی مورد استفاده قرار می‌گیرد. بیشتر مطالعات پیش‌آگهی‌کننده سرطان معده با استفاده از مدل مخاطرات متناسب کاکس تحلیل شده است [۴]. الزامی نبودن یک توزیع احتمالی برای زمان‌های بقا یکی از مزایا و ویژگی‌های بارز مدل مخاطرات متناسب کاکس می‌باشد؛ ولی یک پیش‌فرض مهم و اساسی در این مدل وجود دارد و آن فرض متناسب بودن خطر^۹ برای تمام متغیرهای مستقل موجود در مدل نهایی می‌باشد [۱ و ۲]. در صورت برقراری این فرض، تفسیر مدل به‌دست آمده ساده‌تر و راحت‌تر از مدل‌های پارامتری خواهد بود. اما چنانچه پیش‌فرض‌های مدل‌های پارامتری برقرار باشند، تحلیل قویتری نسبت به مدل کاکس قابل انجام است. قائل شدن برخی مفروضات و انتخاب یک توزیع احتمالی فرضی برای زمان‌های بقا، استنباط آماری را دقیق‌تر نموده و انحراف معیار برآوردها را نسبت به زمانی که چنین مفروضاتی وجود نداشته باشند را کوچک‌تر خواهد کرد [۱ و ۲].

¹ Life Table

² Kaplan Meier Estimator

³ Cox Proportional Hazards Model

⁴ Accelerated Failure Time Model

⁵ Weibull

⁶ Exponential

⁷ Lognormal

⁸ Multiplicative

⁹ Proportional Hazards Assumption

در بررسی مطالعات تحلیل بقاء در خصوص سرطان‌های مختلف، مشخص شده است که تنها ۵ درصد از این مطالعات به بررسی پیش فرض خطرات متناسب پرداخته‌اند [۵]. اگر پیش فرض خطرات متناسب در مدل کاکس برقرار نباشد، نتایج این مدل قابل اعتماد نخواهد بود. همچنین وقتی که توزیع پارامتری خطر مبنا برقرار نباشد، نتایج به دست آمده از مدل‌های پارامتری نیز معتبر نخواهد بود [۶ و ۷]. در چنین وضعیت‌هایی، یک مدل جایگزین، مدل‌های رگرسیونی مخاطرات جمعی می‌باشد. در مدل‌های رگرسیونی مخاطرات جمعی بر خلاف مدل کاکس، متغیرهای مستقل به صورت جمعی^۱ روی میزان خطر مبنا عمل می‌کنند.

۱-۳. مدل‌های رگرسیونی مخاطرات جمعی

در این پژوهش جهت بررسی عوامل مرتبط با بقای بیماران مبتلا به سرطان معده، مدل‌های رگرسیونی مخاطرات جمعی معرفی خواهد شد. مشابه سایر مدل‌های آماری، متغیر زمان لازم تا رخداد پیشامد مورد نظر، T است که توزیع آن به برداری از متغیرهای کمکی وابسته به زمان $\mathbf{Z}(t) = (Z_1(t), \dots, Z_p(t))$ بستگی دارد. فرض می‌شود که میزان مخاطره در زمان t برای فردی با

بردار متغیرهای کمکی $\mathbf{Z}(t)$ ، یک ترکیب خطی از $Z_k(t)$ هاست یعنی

$$h(t|\mathbf{Z}(t)) = \beta_0(t) + \sum_{k=1}^p \beta_k(t) Z_k(t)$$

که در آن $\beta_k(t)$ ها توابعی از متغیرهای کمکی هستند که باید بر اساس داده‌ها برآورد شوند. دو مدل جمعی در این پژوهش ارائه خواهد شد. اولین مدل مربوط به آلن^۲ [۸-۱۰] است که امکان تغییر ضرایب رگرسیونی را در طول زمان فراهم می‌کند. مدل مخاطرات جمعی آلن انعطاف‌پذیری کافی برای ورود اثر متغیرهای کمکی وابسته به زمان را نیز دارد. در این مدل، روش برآورد کمترین مربعات برای برآورد نمودن توابع متغیرهای کمکی $B_k(t) = \int_0^t b_k(u) du$ و خطای معیار این توابع به کار خواهد رفت. این برآوردگرها را می‌توان هموار نمود و برآوردگرهای $b_k(t)$ را به دست آورد (که در این

¹ Additive

² Aalean

پژوهش بررسی نمی‌شود). مدل دوم که توسط لین^۱ و یانگ^۲ [۱۱-۱۳] ارائه شده است، ضرایب وابسته

به زمان رگرسیون $b_k(t)$ را با مقادیر ثابت b_k جایگزین می‌کند:

$$h(t|Z(t)) = b_0(t) + \sum_{k=1}^p b_k Z_k(t)$$

برآوردهای این مدل را می‌توان از طریق معادلات امتیاز^۳ به دست آورد.

۱-۴. مروری بر مطالعات انجام شده در خصوص مدل‌های رگرسیونی مخاطرات جمعی

مبانی نظری مدل‌های رگرسیونی مخاطرات جمعی توسط چندین پژوهشگر از جمله آلن [۸]، برسلو^۴ و دی^۵ [۱۴]، بوکلی^۶ [۱۵]، اوکس^۷ و کاکس^۸ [۱۶]، توماس^۹ [۱۷] و هوفر^{۱۰} و مک کوئیگ^{۱۱} کوئیگ^{۱۱} [۱۸] بررسی شده است. بر خلاف مدل کاکس در این مدل‌ها، متغیرهای مستقل به صورت جمعی روی میزان خطر مبنا عمل می‌کنند. آلن، مدل مخاطرات جمعی ناپارامتری خود و ویژگی‌های آن را در مقالات سال‌های ۱۹۸۰، ۱۹۸۹ و ۱۹۹۳ معرفی کرده است [۸-۱۰]. لین و یانگ در مقالاتی در سال‌های ۱۹۹۴، ۱۹۹۵ و ۱۹۹۷ مدل نیمه پارامتری مخاطرات جمعی را ارائه کردند [۱۱-۱۳]. این مدل می‌تواند به عنوان حالت خاصی از مدل مخاطرات جمعی پارامتری جزئی که توسط مک کوئیگ و ساسینی^{۱۲} ارائه شده است در نظر گرفته شود [۱۹].

¹ Lin

² Ying

³ Score Equation

⁴ Breslow

⁵ Day

⁶ Buckley

⁷ Oakes

⁸ Cox

⁹ Thomas

¹⁰ Huffer

¹¹ McKeague

¹² Sasieni