فصل ۱

مقدمه

برای جواب‌گویی به بسیاری از پرسش‌های علمی، نیاز به گردآوری داده‌ها است. این پرسش‌های علمی در شاخه‌های مختلف علوم به صورت‌های مختلف پیش می‌آیند. به‌عنوان مثال در پزشکی، معنادار بودن تفاوت تأثیر دو دارو بر بیماری خاصی مورد ارزیابی قرار می‌گیرد، در بازاریابی، تأثیر نوع تبلیغ بر تغییر مقدار فروش، مورد کاوش قرار می‌گیرد، و در اقتصاد، تأثیر عوامل مختلف بر بیکاری، مورد بررسی قرار می‌گیرند. در چنین کاربردهایی، نیاز به گردآوری داده‌ها و در نهایت، استخراج اطلاعات است. برای گردآوری داده‌ها آموختن روش‌های مختلف، اجتناب‌ناپذیر است. این روش‌ها در درس نمونه‌گیری به دانشجویان آموخته شده‌اند؛ هرچند ما نیز در این کتاب، اشاره‌ای به برخی روش‌های گردآوری داده‌ها خواهیم داشت. پس از گردآوری داده‌ها، خلاصه‌سازی آن‌ها نیاز به جدول‌بندی دارد، که به چنین جدول‌هایی، جدول‌های پیشایندی گفته می‌شود. پس از جدول‌بندی، استخراج اطلاعات توصیفی از این جدول‌ها مهم است. قسمتی از این کتاب به مباحث توصیفی در مورد داده‌های موجود در جدول‌های پیشایندی اختصاص دارد که داده‌های رسته‌ای نیز نامیده می‌شوند. روش‌های استنباطی که از اوایل قرن بیستم توسط کارل پی‌یرسون و دیگر آمارشناسان معرفی‌شده‌اند نیز در این کتاب، مورد بازبینی قرار می‌گیرند. مباحث پیشرفته‌ی آماری در مورد تحلیل جدول‌های پیشایندی، که از اوایل دهه‌ی ۱۹۶۰ میلادی توسعه پیدا کرده‌اند، قسمت اعظم این کتاب را شامل می‌شوند. با

استفاده از این روش‌ها است که پژوهشگر می‌تواند به بسیاری از پرسش‌های علمی که برخی از آن‌ها در بالا ذکر شده است، پاسخ علمی دهد.

در این فصل، پس از تشریح نوع داده‌ها و مطالعاتی که در این کتاب به تحلیل آن‌ها پرداخته خواهد شد، چندین مثال کاربردی را ارائه می‌کنیم. مثال‌ها به‌گونه‌ای طرح‌ریزی شده‌اند که با مطالب فصل‌های آتی این کتاب مرتبط باشند. در برخی مثال‌ها، آماره‌های توصیفی مهم در تحلیل چندمتغیره‌ی گسسته نیز مورد بحث و بررسی قرار گرفته‌اند. در انتهای فصل نیز برخی دستورهای نرم‌افزار R برای به دست آوردن این آماره‌ها آورده شده‌اند.

۱.۱ مبانی و مفاهیم اولیه

۱.۱.۱ مقیاس‌های اندازه‌گیری و متغیرهای متناظر با آن‌ها

کاملاً روشن است که نوع داده‌ها همچون نوع مسئله‌ای که آمارشناس با آن مواجه است، با تغییر زمینه‌ی تحقیق، تغییر بسیاری می‌یابد. در نتیجه، روش‌هایی که در زمینه‌ای خاص به کار می‌آیند، ممکن است کاربرد اندکی در دیگر زمینه‌ها داشته باشند. داده‌ها ممکن است کمّی یا کیفی باشند. داده‌های کمّی، داده‌هایی هستند که با استفاده از عدد ثبت می‌شوند و دو مقدار مختلف آن‌ها را می‌توان جمع کرد. به‌عنوان مثال، «سن» یک داده‌ی کمّی است. داده‌ی کیفی، داده‌ای است که غیر عددی باشد و توانایی اندازه‌گیری برای آن وجود نداشته باشد و ویژگی آن را نتوان با ارقام ریاضی نمایش داد. برای مثال، رنگ مو، داده‌ای کیفی است. در علوم فیزیکی، داده‌ها لزوماً کمّی با مقیاس‌های دلخواه هستند، در حالی که در علوم اجتماعی و به‌طور کمرنگ‌تر در علوم زیستی، کاربرد داده‌های کیفی بسیار رایج است. این اندازه‌گیری‌های کیفی معمولاً مقادیری را در مجموعه‌ای محدود از رده‌ها اختیار می‌کنند که ممکن است در مقیاس‌های مختلف ثبت شوند.

با توجه به این‌که از این به بعد با واژه‌ی متغیر روبرو می‌شویم، اکنون به معرفی آن می‌پردازیم. منظور ما از متغیر، صفتی است که اندازه‌اش از یک آزمودنی (عضو جامعه یا نمونه‌ی مورد بررسی، که می‌تواند یک انسان، یک گروه یا یک سازمان باشد) به آزمودنی

دیگر تغییر کند. «متغیر» در مقابل «ثابت» است؛ به‌عنوان مثال به هر کتابی «کتاب» گفته می‌شود، ولی به‌عنوان یک متغیر، تعداد صفحات کتاب از یک کتاب به کتاب دیگر تغییر می‌کند.

استیونز (۱۹۵۱، ۱۹۵۹و ۱۹۶۸) به بررسی انواع مقیاس‌ها و ارتباط آن‌ها با روش‌های آماری مورد استفاده پرداخته و چهار مقیاس اسمی، ترتیبی، فاصله‌ای و نسبتی را معرفی کرده است. به‌منظور آشنایی با مقیاس‌های مختلف، به‌اختصار به توصیف آن‌ها می‌پردازیم. متغیرهایی که رده‌های آن‌ها دارای ترتیب نباشند، متغیرهای «اسمی» نامیده می‌شوند. وابستگی دینی و مذهبی (با رده‌های مسلمان، کاتولیک، پروتستان، یهودی و غیره)، نوع موسیقی مورد علاقه و نوع محل سکونت از جمله‌ی متغیرهایی هستند که با مقیاس اسمی اندازه‌گیری می‌شوند.

متغیرهای گسسته‌ی بسیاری دارای رده‌های «ترتیبی» هستند که به چنین متغیرهایی متغیرهای ترتیبی گفته می‌شود. به‌عنوان مثال، اندازه‌ی اتومبیل (با رده‌های بزرگ، متوسط و کوچک)، اثر جانبی استفاده از دارو (با رده‌های بدون اثر جانبی، با اثر جانبی کم، با اثر جانبی زیاد)، و شرایط بیمار از جمله‌ی مثال‌هایی هستند که در آن‌ها متغیرهای اندازه‌گیری‌شده مقیاس ترتیبی دارند. در متغیرهای ترتیبی، فاصله‌ی بین رده‌ها نامعلوم است. متغیرهای «فاصله‌ای» به آن دسته از متغیرها گفته می‌شود که فاصله‌ی عددی بین هر دو مقدار ممکن را نیز اخذ کنند. به‌عنوان مثال، سطح فشار خون، درامد سالانه و طول مدت بیکاری، متغیرهای فاصله‌ای هستند. این‌گونه متغیرها ممکن است که صفر تعریف‌شده نداشته باشند. به‌عنوان مثال، میزان حرارت، متغیری فاصله‌ای است که صفر آن در واحد سانتی‌گراد، مقدار صفر است؛ ولی در مقیاس فارنهایت، ۳۲ است.

متغیرهای «نسبتی» همه‌ی خصوصیات متغیرهای فاصله‌ای را دارند به‌علاوه‌ی این‌که یک نقطه‌ی صفر تعریف‌شده نیز دارند. به‌عنوان مثال می‌توان به درامد افراد اشاره کرد که در هر واحدی اندازه‌گیری شود صفر آن، صفر معمولی است.

روش اندازه‌گیری متغیرها نوع رده‌بندی آن‌ها را مشخص می‌کند. به‌عنوان مثال، تحصیلات در صورتی یک متغیر اسمی است که با دو رده با عنوان‌های «تحصیلات در مدارس عمومی» و «تحصیلات در مدارس خصوصی» رده‌بندی شود. همین متغیر، ترتیبی است هرگاه بر اساس بالاترین درجه‌ی کسب‌شده (مثلاً به‌صورت بی‌سواد، زیر دیپلم، دیپلم،

فوق دیپلم، لیسانس، فوق لیسانس و دکتری) اندازه‌گیری شود.

رده‌بندی چهار مقیاس اشاره شده بر اساس کمّی و کیفی بودن به‌صورت زیر بیان می‌شود. متغیرهای اسمی همواره کیفی هستند، چرا که در آن‌ها تفاوت رده‌ها از لحاظ کیفیت است نه از لحاظ کمّیت. متغیرهای فاصله‌ای کمّی هستند، چرا که تفاوت سطوح مجزا، از لحاظ مقدار مشخصه‌ی مورد نظر است. متغیرهای نسبتی نیز مانند متغیرهای فاصله‌ای، کمّی هستند. رده‌بندی متغیرهای ترتیبی بر حسب کمّی یا کیفی بودن، مبهم است. در بعضی موارد، آن‌ها به‌عنوان متغیرهای کیفی در نظر گرفته می‌شوند و از روش‌هایی که برای متغیرهای اسمی به کار می‌روند، برای تحلیل آن‌ها استفاده می‌شود؛ اما در بسیاری موارد، متغیرهای ترتیبی بیش‌تر به متغیرهای فاصله‌ای شباهت دارند تا به متغیرهای اسمی، که در این موارد از تحلیل‌های کمّی استفاده می‌شود. حالتی را که در آن، داده‌ها کیفی باشند و از مقیاس ترتیبی برخوردار باشند، می‌توان یکی از رایج‌ترین حالت‌ها در زمینه‌های اجتماعی برشمرد، که با در نظر نگرفتن این ترتیب، بخشی از اطلاعات موجود در داده‌ها از دست می‌رود.

در این کتاب با متغیرهایی سر و کار داریم که در مقیاس‌های ترتیبی و اسمی اندازه‌گیری می‌شوند، یا اگر در مقیاس‌های فاصله‌ای و نسبتی اندازه‌گیری شده‌اند، آن‌ها را با رده‌بندی به متغیرهای اسمی یا ترتیبی تبدیل کرده‌اند و اطلاعی از اندازه‌گیری اولیه‌ی آن‌ها در مقیاس فاصله‌ای یا نسبتی در دست نیست. در بخش‌های بعدی به‌تفصیل به تحلیل متغیرهای ترتیبی و اسمی پرداخته می‌شود.

متغیرها را بر حسب تأثیر و تأثر یا وابستگی نیز به دو دسته تقسیم می‌کنند: متغیرهای وابسته و متغیرهای مستقل. متغیر وابسته (پاسخ) متغیری است که تحت تأثیر متغیرهای دیگر است و معمولاً آثار سایر متغیرها بر روی آن مورد بررسی قرار می‌گیرد. متغیر مستقل، متغیری است که عامل مؤثر (اثرگذار) بر متغیر وابسته‌ی مورد نظر است. به‌عنوان مثال، سواد مادر به‌عنوان یک متغیر مستقل بر پیشرفت تحصیلی فرزند او (به‌عنوان یک متغیر وابسته) تأثیر می‌گذارد.

۲.۱.۱ انواع مطالعه‌ها و روش‌های گردآوری داده‌ها

مسئله‌ی مهم دیگر در اجرای یک تحقیق، نوع مطالعه و روش گردآوری داده‌ها است. آن دسته از مطالعات را که در آن‌ها فقط یک متغیر پاسخ در مقطعی معلوم از زمان مورد بررسی است، مطالعات مقطعی می‌نامند. مطالعاتی نظیر رگرسیون، تحلیل واریانس و تحلیل کوواریانس از آن جمله‌اند.

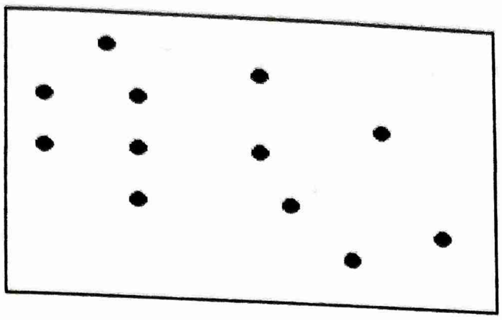
به مطالعاتی که در آن‌ها متغیر پاسخ برای آزمودنی‌ها چندین‌بار در طول زمان اندازه‌گیری می‌شود، مطالعات طولی می‌گویند.

در این میان، مطالعات مقطعی هرچند از لحاظ قابلیت اجرا و تفسیر، آسان‌تر به نظر می‌رسند، هیچ‌گونه اطلاعی در مورد چگونگی تغییر تأثیر عوامل بر پاسخ مورد بررسی در طول زمان به دست نمی‌دهند. به همین دلیل، مطالعات طولی که با بهره‌گیری از اندازه‌گیری‌های تکراری، امکان مطالعه‌ی تغییرات درون هر واحد را نیز میسر می‌سازند، در مطالعات اجتماعی و پزشکی، بیش‌تر مورد علاقه هستند تا در مطالعات مقطعی.

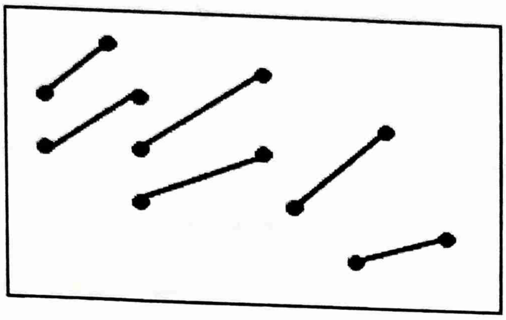
به مثال زیر از دیگل و دیگران (۲۰۰۲) توجه کنید. در شکل ۱.۱ قدرت خواندن در مقابل سن برای کودکان در یک مطالعه‌ی مقطعی رسم شده است که نشان می‌دهد قدرت خواندن در میان کودکان با افزایش سن کاهش می‌یابد. در شکل ۲.۱ فرض می‌کنیم داده‌های مشابهی در یک مطالعه‌ی طولی به دست آمده‌اند که در آن هر فرد، دو بار اندازه‌گیری شده است. در این شکل، واضح است که در حالی که توانایی خواندن در کودکان جوان‌تر، سطح بالاتری دارد، همه‌ی آزمودنی‌ها توانایی خود را با گذشت زمان، بهبود می‌دهند.

نکته‌ی قابل توجه در این مثال، این است که مطالعه‌ی طولی (شکل ۲.۱) می‌تواند اثر زمان برای هر آزمودنی (اثر سن) را از اثر گروه که تفاوت‌های آزمودنی‌ها در سطح پایه‌ی آن‌ها است جدا کند. در مثال بالا اثر سن نشان می‌داد که همه‌ی کودکان، توانایی خواندن‌شان را افزایش می‌دهند و اثر گروه نشان می‌دهد که توانایی خواندن کودکان کم‌سن‌تر، بیش‌تر از کودکان بزرگ‌تر است؛ در حالی که مطالعات مقطعی فقط اثر گروه را بیان می‌کنند.

قبل از معرفی داده‌های مربوط به مطالعات مقطعی و طولی گسسته که در فصل‌های بعدی لازم است از آن‌ها استفاده شود، بهتر است به معرفی روش‌های مختلف گردآوری



شکل ۱.۱: توانایی خواندن کودکان در مطالعات مقطعی



شکل ۲.۱: توانایی خواندن کودکان در مطالعات طولی

داده‌ها بپردازیم. مطالعات و داده‌هایی که از اطلاعات گذشته‌ی افراد استفاده می‌کنند، به مطالعات گذشته‌نگر معروف‌اند. این‌گونه مطالعات به مطالعات مورد– شاهد نیز معروف‌اند. بر عکس، مطالعاتی که در آن‌ها آزمودنی‌های نمونه در زمان خاصی مورد مطالعه قرار گیرند و همچنین وضعیت آن‌ها پس از گذشت چندین سال مجدداً مورد بررسی قرار گیرد، مطالعات آینده‌نگر نامیده می‌شوند. دو نوع مطالعه‌ی آینده‌نگر وجود دارد: آزمایش‌های بالینی و مطالعات هم‌گروهی. در آزمایش‌های بالینی، آزمودنی‌های مورد مطالعه به‌صورت تصادفی به گروه‌های مختلف (به‌عنوان مثال، سیگاری و غیر سیگاری در مطالعه‌ی سرطان ریه) تخصیص می‌یابند و وضعیت ریه‌ی آن‌ها پس از مدتی مورد بررسی قرار می‌گیرد. در مطالعات هم‌گروهی، آزمودنی‌ها انتخاب خودشان را برای این‌که در چه گروهی باشند بیان می‌کنند. روش دیگر گردآوری داده‌ها، مطالعات مقطعی است که در آن‌ها آزمودنی‌ها پس از نمونه‌گیری، به‌طور همزمان برای دو یا چند متغیر طبقه‌بندی می‌شوند.

مطالعات مورد– شاهد، هم‌گروهی و مقطعی، مطالعات «مشاهده‌ای» نامیده می‌شوند. بر عکس، آزمایش بالینی، یک مطالعه‌ی «آزمایشی» است که در آن، محقق روی

آزمودنی‌هایی که درمان را می‌پذیرند، کنترل دارد. از این جنبه، مطالعات آزمایشی عالی‌ترین شکل پژوهش‌اند و دقیق‌ترین شرایط را برای آزمون فرضیه‌های علّی فراهم می‌سازند.

۳.۱.۱ دورنمای کتاب

در این کتاب به معرفی مطالعات مقطعی و طولی و مدل‌های مورد استفاده برای تحلیل داده‌های رده‌بندی‌شده پرداخته می‌شود. در ادامه‌ی این فصل به معرفی برخی کاربردهای واقعی و داده‌های منتج از آن‌ها پرداخته خواهد شد که در فصل‌های بعدی از آن‌ها استفاده می‌شود. در این مثال‌ها به‌ترتیب، کاربردهایی با پاسخ‌های دودویی، ترتیبی و اسمی در مطالعات مقطعی و طولی تشریح می‌شوند. در فصل دوم به بررسی مطالعات مقطعی برای پاسخ‌های دودویی خواهیم پرداخت و مدل‌های مختلف برای تحلیل داده‌های مقطعی دودویی و روش‌های محاسباتی برای براورد کردن پارامترها معرفی خواهند شد. همچنین در این فصل، مثال‌هایی برای پاسخ‌های مقطعی دودویی بیان می‌شود. همین مباحث نیز به‌ترتیب برای پاسخ‌های ترتیبی و اسمی در مطالعات مقطعی، به‌ترتیب در فصل‌های سوم و چهارم بیان می‌شوند. معرفی مدل‌های لگ‌خطی برای جدول‌های پیشایندی از جمله‌ی مباحث مطرح در فصل پنجم است. بررسی مطالعات طولی با پاسخ‌های دودویی، همراه با مدل‌های مختلف و روش‌های مختلف براورد کردن پارامترها در فصل ششم بیان شده است و در پایان این فصل نیز در چند مثال کاربردی به تحلیل این نوع داده‌ها پرداخته می‌شود. همین مباحث برای مطالعات طولی با پاسخ ترتیبی و اسمی، به‌ترتیب در فصل‌های هفتم و هشتم مورد بررسی قرار می‌گیرند.

۲.۱ مثال‌های کاربردی

به‌منظور ایجاد انگیزه برای رشد و توسعه‌ی مبانی نظری و همچنین روش‌شناسی، به معرفی بعضی داده‌های واقعی می‌پردازیم. هرچند تحلیل آماری برخی از این داده‌ها در فصل‌های بعدی صورت می‌گیرد، در این‌جا برخی آماره‌های توصیفی و نمایش‌های

گرافیکی برای تشریح اولیه‌ی داده‌ها ارائه می‌شوند.

۱.۲.۱ داده‌های مقطعی با پاسخ دودویی

در این زیربخش، چهار کاربرد واقعی در مطالعات مقطعی برای پاسخ‌های دودویی آورده می‌شود.

۱.۱.۲.۱ بررسی اثر دوزهای متفاوت دیسولفید کربن بر روی سوسک‌ها

مثالی از پاسخ‌های دودویی مقطعی، بررسی اثر دوزهای متفاوت دیسولفید کربن بر روی سوسک‌ها است. این مثال، یکی از پرکاربردترین مثال‌ها در زمینه‌ی سم‌شناسی و داروسازی است. متغیر پاسخ در این گونه مثال‌ها تنها دارای دو حالت «زنده ماندن» و «زنده نماندن» است که نوع پاسخ در آن (دودویی با دو حالت شکست و موفقیت) از موارد بحث در این کتاب است.

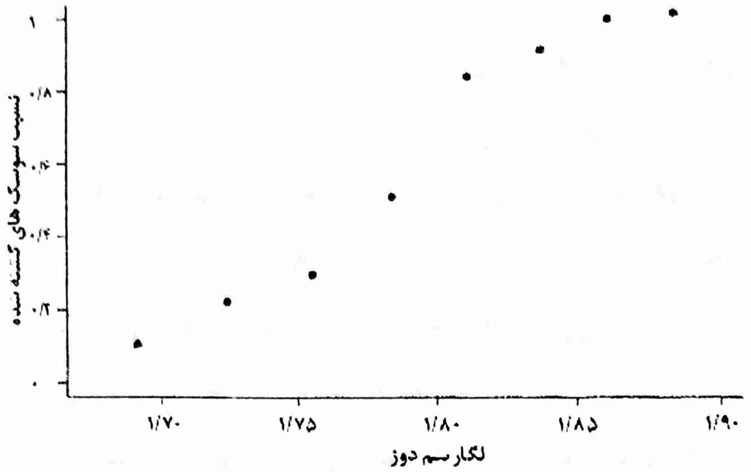
داده‌های جدول ۱.۱ که منبع آن بلیس (۱۹۳۵) است و اگرستی (۲۰۰۲، ص. ۲۴۷) نیز از آن استفاده کرده است، تعداد سوسک‌های کشته‌شده را نشان می‌دهد که بعد از ۵ ساعت قرار گرفتن سوسک‌ها در معرض گاز دیسولفید کربن در غلظت‌های متفاوت اندازه‌گیری شده است. غلظت به‌صورت لگاریتم دوز بیان شده است.

جدول ۱.۱: اثر دوزهای مختلف دیسولفید کربن بر روی سوسک‌ها (منبع: بلیس، ۱۹۳۵)

لگاریتم دوز	۱/۶۹۱	۱/۷۲۴	۱/۷۵۵	۱/۷۸۴	۱/۸۱۱	۱/۸۳۷	۱/۸۶۱	۱/۸۸۴
تعداد سوسک‌ها	۵۹	۶۰	۶۲	۵۶	۶۳	۵۹	۶۲	۶۰
تعداد کشته‌شده‌ها	۶	۱۳	۱۸	۲۸	۵۲	۵۳	۶۱	۶۰
نسبت کشته‌شده‌ها	۰/۱۰۲	۰/۲۱۷	۰/۲۹۰	۰/۵۰۰	۰/۸۲۵	۰/۸۹۸	۰/۹۸۴	۱/۰۰

همان‌گونه که در شکل ۳.۱ ملاحظه می‌شود، با افزایش لگاریتم دوزها نسبت سوسک‌های کشته‌شده افزایش می‌یابد. البته افزایش نسبت سوسک‌های کشته‌شده رابطه‌ای خطی با افزایش لگاریتم دوزها ندارد. ملاحظه می‌شود که وقتی لگاریتم دوزهای دیسولفید کربن از ۱/۷۸۴ به ۱/۸۱۱ و بیش‌تر تبدیل می‌شود، نسبت سوسک‌های کشته‌شده افزایش چشمگیری داشته است. تحلیل و بررسی کامل این مثال در فصل بعدی که مبحث مربوط به مطالعات مقطعی با پاسخ‌های دودویی است، ارائه خواهد شد.

۲.۱.۲.۱ بررسی اثر پرتودهی گاز دی‌اکسید نیتروژن و مدت زمان آن بر احتمال مرگ



شکل ۳.۱: نسبت سوسک‌های کشته‌شده بر اثر لگاریتم دوزهای مختلف دیسولفید کربن

موش‌ها

داده‌های جدول ۲.۱ که توسط لارسن و دیگران (۱۹۷۹) گزارش شده است، مربوط به تعدادی موش تحت آزمایش می‌باشد که این موش‌ها به m گروه تعلق دارند. فرض کنید $y_۱،\dots،y_m$ فراوانی‌های مشاهده‌شده‌ی m توزیع دوجمله‌ای مستقل از هم با احتمال‌های $p_۱،\dots،p_m$ و اندازه‌های نمونه‌ی $n_۱،\dots،n_m$ باشند به‌طوری که

$$y_i \sim \mathrm{Bin}(\mathrm{n_i}, \mathrm{p_i}), \qquad i = ۱, \dots, m.$$

جدول ۲.۱: موش‌های در معرض پرتودهی $NO_۲$ (منبع: لارسن و دیگران، ۱۹۷۹)

شماره‌ی گروه	درجه‌ی پرتودهی $NO_۲$	مدت زمان پرتودهی	تعداد موش‌های مرده (y_i)	تعداد موش‌ها (n_i)
۱	۱/۵	۹۶	۴۴	۱۲۰
۲	۱/۵	۱۶۸	۳۷	۸۰
۳	۱/۵	۳۳۶	۴۳	۸۰
۴	۱/۵	۵۰۴	۳۵	۶۰
۵	۳/۵	۰/۵	۲۹	۱۰۰
۶	۳/۵	۱	۵۳	۲۰۰
۷	۳/۵	۲	۱۳	۴۰
۸	۳/۵	۳	۷۵	۲۰۰
۹	۳/۵	۵	۲۳	۴۰
۱۰	۳/۵	۷	۱۵۲	۲۸۰
۱۱	۳/۵	۱۴	۵۵	۸۰
۱۲	۳/۵	۲۴	۹۸	۱۴۰
۱۳	۳/۵	۴۸	۱۲۱	۱۶۰
۱۴	۷	۰/۵	۵۲	۱۲۰
۱۵	۷	۱	۶۲	۱۲۰
۱۶	۷	۱/۵	۶۱	۱۲۰
۱۷	۷	۲	۸۶	۱۲۰

در داده‌های جدول ۲.۱، y_i معرف تعداد موش‌های مرده در iامین گروه $(i = ۱, \ldots, ۱۷)$ است و n_i معرف تعداد موش‌های مورد آزمایش در گروه iام است که در معرض گاز دی‌اکسید نیتروژن قرار گرفته‌اند. گیریم فرض‌های رایج در مورد توزیع دوجمله‌ای، برقرار باشند و p_i معرف احتمال مرگ برای هر موش در گروه iام باشد. مسئله‌ی مورد بررسی این است که آیا احتمال مرگ به دو متغیر کمکی «درجه‌ی پرتودهی گاز دی‌اکسید نیتروژن» و «مدت زمان پرتودهی» بستگی دارد یا نه.

۳.۱.۲.۱ درمان سنگ کلیه (پارادوکس سیمپسون)

در مثال زیر به اهمیتِ در نظر گرفتن همه‌ی اطلاعات کمکی در تعیین احتمال موفقیت سنگ کلیه و به پارادوکسی اشاره می‌شود که به پارادوکس سیمپسون (سیمپسون، ۱۹۵۱) معروف است. جدول ۳.۱ را در نظر بگیرید که در آن، سنگ کلیه تحت درمان (یا تیمار) A (شیوه‌ی باز) یا تحت درمان B (شیوه‌ی پوستی) قرار گرفته است.

جدول ۳.۱: درمان سنگ کلیه تحت دو تیمار متفاوت (منبع: چَریگ و دیگران، ۱۹۸۶)

تیمار	درمان سنگ کلیه		کل
	موفقیت	شکست	
تیمار A	۲۷۳	۷۷	۳۵۰
تیمار B	۲۸۹	۶۱	۳۵۰

جدول ۴.۱: درمان سنگ کلیه تحت دو تیمار متفاوت به تفکیک اندازه‌ی سنگ کلیه

تیمار	درمان سنگ کلیه (اندازه‌ی سنگ: بزرگ)		کل
	موفقیت	شکست	
تیمار A	۱۹۲	۷۱	۲۶۳
تیمار B	۵۵	۲۵	۸۰
تیمار	درمان سنگ کلیه (اندازه‌ی سنگ: کوچک)		کل
	موفقیت	شکست	
تیمار A	۸۱	۶	۸۷
تیمار B	۲۳۴	۳۶	۲۷۰

شانس موفقیت با استفاده از تیمار A برابر با $\frac{۲۷۳}{۳۵۰} = ۰/۷۸$ است و شانس موفقیت با استفاده از تیمار B برابر با $\frac{۲۸۹}{۳۵۰} = ۰/۸۳$ است. با مقایسه‌ی این دو نسبت، به نظر می‌رسد که تیمار B مؤثرتر است (البته معنادار بودن تفاوت باید با در نظر گرفتن واریانس براوردگرها یا با استفاده از یک مدل آماری بررسی شود). این نشان می‌دهد که شانس موفقیت به

شکست استفاده از تیمار A برابر با $\frac{۰/۷۸}{۰/۲۲} = ۳/۵۴$ است، ولی شانس موفقیت به شکست استفاده از تیمار B برابر با $\frac{۰/۸۳}{۰/۱۷} = ۴/۸۸$ است (به این شانس‌ها یا بخت‌ها، «بخت حاشیه‌ای» گفته می‌شود). این‌ها نشان می‌دهند که بخت موفقیت استفاده از تیمار B، تقریباً $\frac{\frac{۰/۸۳}{۰/۱۷}}{\frac{۰/۷۸}{۰/۲۲}} = ۱/۳۷۸$ برابر بخت موفقیت استفاده از تیمار A است.
اگر اطلاعات اندازه‌ی سنگ کلیه را نیز به اطلاعات بالا اضافه کنیم، اطلاعات جدول ۴.۱ را خواهیم داشت.

اکنون زمانی که اندازه‌ی سنگ بزرگ است، شانس موفقیت با تیمار A برابر با $\frac{۱۹۲}{۲۶۳} = ۰/۷۳$ و شانس موفقیت با تیمار B برابر با $\frac{۵۵}{۸۰} = ۰/۶۹$ است که به نظر می‌آید شانس موفقیت با تیمار A بیش‌تر است. زمانی که اندازه‌ی سنگ کوچک است، شانس موفقیت با تیمار A برابر با $\frac{۸۱}{۸۷} = ۰/۹۳$ و شانس موفقیت با تیمار B برابر با $\frac{۲۳۴}{۲۷۰} = ۰/۸۷$ است که دوباره به نظر می‌آید شانس موفقیت با تیمار A بیش‌تر است. این نتیجه‌گیری دوگانه را که با استفاده از جدول ۳.۱ (جدول حاشیه‌ای) عملکرد تیمار B بهتر به نظر می‌آید و با استفاده از جدول ۴.۱ عملکرد تیمار A بهتر است، پارادوکس سیمپسون می‌نامند. اگر بخت شرطیِ موفقیت تیمار B را به‌صورت بخت موفقیت با تیمار B در سطح داده‌شده‌ی متغیر «اندازه‌ی سنگ کلیه» تعریف کنیم، بخت موفقیت شرطی با تیمار B، وقتی اندازه‌ی سنگ کلیه بزرگ است، برابر با $\frac{۰/۶۹}{۰/۳۱} = ۲/۲۲۶$ است. همین بخت برای تیمار A عبارت است از $\frac{۰/۷۳}{۰/۲۷} = ۲/۷۰۴$ که نشان می‌دهد بخت شرطی موفقیت تیمار B، تقریباً $\frac{\frac{۰/۶۹}{۰/۳۱}}{\frac{۰/۷۳}{۰/۲۷}} = ۰/۸۲۳$ برابر بخت شرطی موفقیت با تیمار A در زمانی است که اندازه‌ی سنگ کلیه بزرگ باشد. نسبت دو بخت بالا را «نسبت بخت‌های شرطی» می‌نامند.
در این مثال، متغیر ایجادکننده‌ی اختلال (اندازه‌ی سنگ کلیه)، قبل از هر گونه نتیجه‌گیری باید کنترل شود. در یک مدل لوجیت برای پاسخ‌های دودویی خواهیم دید که چگونه می‌توان اثر دو متغیر اندازه‌ی سنگ کلیه و نوع تیمار (به‌عنوان دو متغیر دودویی تبیینی) بر بخت موفقیت را مورد کاوش قرار داد.

اکنون به معرفی آماره‌ی پی‌یرسون برای آزمون دوجمله‌ای‌های مستقل با احتمال‌های مفروض در مثال درمان سنگ کلیه می‌پردازیم. فرض کنید در جدول ۳.۱ متغیر Y را برای موفقیت یا شکست درمان (۱: موفقیت و ۲: شکست) و متغیر X را

برای نوع تیمار (۱: تیمار A و ۲: تیمار B) به کار ببریم. به دنبال آن هستیم که بر اساس اطلاعات این جدول $H_0 : \Pr(Y = 1|X = 1) = \Pr(Y = 1|X = 2)$ را در مقابل $H_1 : \Pr(Y = 1|X = 1) \neq \Pr(Y = 1|X = 2)$ آزمون کنیم. جدول احتمال‌های شرطی زیر را در نظر بگیرید:

X	$Y = 1$	$Y = 2$
$X = 1$	$\pi_{1\|1}$	$\pi_{2\|1}$
$X = 2$	$\pi_{1\|2}$	$\pi_{2\|2}$

که در آن $\pi_{j|i} = \Pr(Y = j|X = i)$ و آزمون فرض بالا معادل است با آزمون $H_0 : \pi_{1|1} = \pi_{1|2}$ در مقابل $H_1 : \pi_{1|1} \neq \pi_{1|2}$. اگر مقادیر نمونه در همه‌ی خانه‌های جدول، به اندازه‌ی کافی بزرگ باشند، می‌توان از آماره‌ی خی‌دو به‌صورت

$$\chi^2 = \sum_j \sum_i \frac{(n_{ij} - \hat{\mu}_{ij})^2}{\hat{\mu}_{ij}}$$

استفاده کرد که در آن، n_{ij} فراوانی مشاهده‌شده در سطر iام و ستون jام جدول است و $\hat{\mu}_{ij} = n\hat{\pi}_{i+}\hat{\pi}_{+j} = \frac{n_{i+}n_{+j}}{n}$. در رابطه‌ی اخیر داریم $\pi_{i+} = \sum_j \pi_{ij}$ و $\pi_{+j} = \sum_i \pi_{ij}$ و $n_{i+} = \sum_j n_{ij}$ و $n_{+j} = \sum_i n_{ij}$. این آماره تحت H_0 دارای توزیع خی‌دو با $(I - 1)(J - 1) = (2 - 1)(2 - 1) = 1$ درجه‌ی آزادی است که در آن I تعداد سطرها و J تعداد ستون‌های جدول است (مود و دیگران، ۱۹۷۴، فصل ۹ را ببینید).

۴.۱.۲.۱ آزمون دقیق فیشر در مثال تأثیر دود تنباکو بر ایجاد تومور

آماره‌های نیکویی برازش، همچون آماره‌ی χ^2 پی‌یرسون و آماره‌ی نسبت درست‌نمایی تعمیم‌یافته، وقتی اندازه‌ی نمونه کوچک است، مناسب نیستند. برخی آمارشناسان فقط در حالتی که فراوانی همه‌ی خانه‌های جدول بیش‌تر از ۵ باشد، از آماره‌های مذکور استفاده می‌کنند (لی، ۱۹۹۸، ص. ۳۲). برای جدول‌های پیشایندی ۲ × ۲ با نمونه‌ی کوچک و حاشیه‌های سطری و ستونی داده‌شده (مانند جدول ۵.۱) فقط تعداد مشاهدات در یکی از خانه‌ها (به‌عنوان مثال، خانه‌ی سطر اول و ستون اول)، یک متغیر تصادفی است. اگر این متغیر تصادفی را N_{11} بنامیم که در جدول، مقدار n_{11} را اخذ می‌کند، می‌توانیم از آزمون دقیق فیشر برای آزمون استقلال دو متغیر تصادفی استفاده کنیم.

جدول ۵.۱: نمایش جدول پیشایندی ۲ × ۲

X	Y: ۱	Y: ۲	
۱	n_{11}	n_{12}	n_{1+}
۲	n_{21}	n_{22}	n_{2+}
	n_{+1}	n_{+2}	n

در جدول پیشایندی ۲ × ۲، استقلال برقرار خواهد بود اگر به‌ازای هر i و j داشته باشیم $\pi_{ij} = \pi_{i+}\pi_{+j}$، که در آن $\pi_{ij} = \Pr(X = i, Y = j)$، $\pi_{i+} = \Pr(X = i)$ و $\pi_{+j} = \Pr(Y = j)$. بخت مقدار ۱ برای Y در سطح ۱ برای X عبارت است از

$$\frac{\pi_{1|1}}{\pi_{2|1}} = \frac{\frac{\pi_{11}}{\pi_{1+}}}{\frac{\pi_{12}}{\pi_{1+}}} = \frac{\pi_{11}}{\pi_{12}},$$

که در آن $\pi_{j|i} = \Pr(Y = j|X = i)$ است و بخت مقدار ۱ برای Y در سطح ۲ برای X عبارت است از

$$\frac{\pi_{1|2}}{\pi_{2|2}} = \frac{\frac{\pi_{21}}{\pi_{2+}}}{\frac{\pi_{22}}{\pi_{2+}}} = \frac{\pi_{21}}{\pi_{22}}.$$

در صورت استقلال، نسبت بخت‌ها باید ۱ باشد و خواهیم داشت

$$\theta = \frac{\frac{\pi_{11}}{\pi_{12}}}{\frac{\pi_{21}}{\pi_{22}}} = 1,$$

که نتیجه می‌دهد ۱ = $\theta = \frac{\pi_{11}\pi_{22}}{\pi_{21}\pi_{12}}$. برای آزمون ۱ = $H_0 : \theta$ در مقابل همبستگی مثبت $(H_1 : \theta > 1)$، از آن‌جا که N_{11} (تحت H_0) دارای توزیع فوق‌هندسی زیر است

$$\Pr(N_{11} = n_{11}|\theta = 1) = \frac{\binom{n_{1+}}{n_{11}}\binom{n - n_{1+}}{n_{+1} - n_{11}}}{\binom{n}{n_{+1}}},$$

که در آن $n_{11}\epsilon\{\max(0, n_{1+} + n_{+1} - n), \min(n_{1+}, n_{+1})\}$، می‌توان از P−مقدار برای آزمون فرض مطلوب به‌صورت زیر استفاده کرد:

$$P\text{−مقدار} = \Pr_{H_0}(N_{11} \geq n_{11(\text{obs})}) = \sum_{\{n_{11}:N_{11}\geq n_{11(\text{obs})}\}} \Pr(N_{11} = n_{11}|\theta = 1)$$

که در آن $n_{11(\text{obs})}$ مقدار مشاهده‌شده در خانه‌ی با سطر و ستون ۱ است. توجه داشته باشید که یک مقدار بزرگ N_{11}، موافق H_1 است.

اکنون موارد فوق را به‌عنوان مثال در داده‌های تأثیر دود تنباکو بر مخاطره داشتن تومور که در زیر آمده است، مورد ارزیابی قرار می‌دهیم. بر اساس داده‌های جدول ۶.۱ که در آن تأثیر دود تنباکو بر مخاطره داشتن تومور مورد ارزیابی قرار گرفته است (اسنبرگ، ۱۹۵۲)، از ۲۳ موشی که در معرض دود تنباکو قرار گرفته‌اند، ۲۱ موش تومور داشته‌اند و از ۳۲ موشی که در معرض دود قرار نگرفته‌اند، ۱۹ موش تومور داشته‌اند. سؤال مورد نظر این است که آیا می‌توان پذیرفت که دود تنباکو و داشتن تومور، پیوندی ندارند؟ فرض عدم پیوند را در مقابل این‌که دود تنباکو موجب می‌شود موش‌های بیش‌تری تومور داشته باشند (پیوند وجود داشته باشد)، آزمون می‌کنیم.

جدول ۶.۱: بررسی تأثیر دود تنباکو بر ایجاد تومور (منبع: اسنبرگ، ۱۹۵۲)

	داشتن تومور	نداشتن تومور	
در معرض دود قرار گرفتن	۲۱	۲	۲۳
کنترل	۱۹	۱۳	۳۲
کل	۴۰	۱۵	۵۵

چون آزمون $H_{\circ} : \theta = ۱$ در مقابل $H_{۱} : \theta > ۱$ مد نظر است، داریم:

$$\text{مقدار}-P = \Pr_{H_{\circ}}(N_{۱۱} \geq ۲۱) = ۰/۰۰۸۳.$$

P-مقدار به دست آمده، فرض صفر را به‌طور قوی رد می‌کند (برای محاسبه‌ی این P-مقدار می‌توانید به نشانی اینترنتی $http://www.langsrud.com/fisher.html$ مراجعه کنید).

برخی پیشنهاد کرده‌اند (لانکاستر، ۱۹۶۱؛ اگرستی، ۲۰۰۲) که وقتی فراوانی برخی خانه‌های جدول کوچک است، برای استنباط از مقدار mid$-P$ استفاده شود. برای آماره‌ی آزمون T با مقدار مشاهده‌شده‌ی $t_{\circ}$ و آزمون یک‌طرفه‌ای که مقادیر بزرگ T منجر به رد $H_{\circ}$ می‌شود، داریم:

$$\text{مقدار}-(\text{mid}-P) = \tfrac{۱}{۲}\Pr_{H_{\circ}}(T = t_{\circ}) + \Pr_{H_{\circ}}(T > t_{\circ}).$$

بنا بر این $(\text{mid}-P)$-مقدار کمتر از P-مقدار است و در مورد رد $H_{\circ}$، کمتر محافظه‌کارانه عمل می‌کند. در مورد مثال جدول ۶.۱ داریم:

$$\text{مقدار}-(\text{mid}-P) = \tfrac{۱}{۲}\Pr_{H_{\circ}}(N_{۱۱} = ۲۱) + \Pr_{H_{\circ}}(N_{۱۱} > ۲۱),$$

که احتمال‌های فوق از توزیع فوق‌هندسی قابل محاسبه‌اند.

۱.۲.۲ داده‌های مقطعی با پاسخ ترتیبی

در این زیربخش، دو کاربرد واقعی در مطالعات مقطعی برای پاسخ‌های ترتیبی با استفاده از طرح آمارگیری از هزینه و درامد خانوارها و طرح آمارگیری نیروی کار مرکز آمار ایران آورده می‌شود.

۱.۱.۲.۲ مقایسه‌ی فقر با استفاده از داده‌های طرح آمارگیری از هزینه و درامد خانوارها در سال ۱۳۸۴

مثالی از داده‌های تبدیل‌شده به داده‌های مقطعی ترتیبی، داده‌های طرح آمارگیری از هزینه و درامد خانوارها است که توسط مرکز آمار ایران گردآوری می‌شود. در این مثال از داده‌های ثبت‌شده برای هر خانوار شهری استفاده می‌کنیم، که شامل درامد و هزینه‌ی خانوار[۱] (HIE) می‌باشد. در این مثال که بر اساس اطلاعات سال ۱۳۸۴ است، از نمونه‌گیری‌های تصادفی پیچیده با حجم زیاد استفاده شده است که طی آن، مقدار زیادی اطلاعات از خانوارهای مورد بررسی به دست می‌آید. در کل، ۱۱۶۱۹ فرد شهری سرپرست خانوار، مورد بررسی قرار گرفته‌اند. مهم‌ترین دلیل سنجش فقر، ضرورتِ دست‌یابی به یک عدد صِرف برای مکان و تاریخی خاص نیست، بلکه مقایسه‌ی فقر است. مقایسه‌ی فقر، ارزیابی این نکته است که از میان دو وضعیت، کدام‌یک با فقر بیش‌تری توأم است. مقایسه‌های عددی فقر، مستلزم آن است که اندازه‌ی تفاوت موجود در میزان فقر به‌طور کمّی بیان شود. اما زمانی که هدف، مقایسه در مقوله‌ی فقر باشد، ارزیابی ترتیبی ضرورت دارد (راوالیون، ۲۰۰۴). از طرف دیگر، اغلب داده‌هایی که در تحلیل فقر به کار می‌روند دارای خطای بسیار می‌باشند و گریز از این امر، اجتناب‌ناپذیر است. پس رده‌بندی متغیر هزینه‌ی کل بر اساس خطوط فقر، رهنمون تحلیلی دقیق‌تر در مقوله‌ی فقر خواهد بود.

خط فقر مورد استفاده در این بررسی، خط فقر مطلق بر مبنای ۲۳۰۰ کالری و ۳۹۰۰۵۸ ریال به‌عنوان خط فقر در سال ۱۳۸۳ است (باقری و دیگران، ۱۳۸۳) که با توجه به نرخ تورم ۱۴ درصدی در سال ۱۳۸۴، خط فقر در سال ۱۳۸۴ به ۴۴۴۶۶۶/۱ ریال بالغ می‌شود. برای بررسی فقر به‌عنوان تابعی از هزینه‌ی کل، درصد افراد در رده‌های مختلف

---

[۱] Household Income and Expenditure

وضعیت فقر در سال ۱۳۸۴ را طبق جدول ۷.۱ معرفی می‌کنیم. در این جدول، در صورتی که متغیر مورد علاقه‌ی وضعیت فقر فرد (Y) برابر با ۱ باشد، فرد را فقیر می‌نامیم؛ در صورتی که $Y = ۲$ باشد، فرد را شبه‌فقیر گوییم؛ در صورتی که $Y = ۳$ باشد، فرد را غیر فقیر می‌نامیم، و در صورتی که $Y = ۴$ باشد، فرد را شبه‌غنی یا غنی قلمداد می‌کنیم.

جدول ۷.۱: درصد افراد در رده‌های مختلف وضعیت فقر در سال ۱۳۸۴

متغیر پاسخ (Y)	فراوانی	درصد
هزینه‌ی ماهانه‌ی هر فرد خانوار، کمتر یا مساوی ۴۴۴۶۶۶/۱ ریال است ($Y = ۱$)	۲۱۱۸	۱۸/۲
هزینه‌ی ماهانه‌ی هر فرد خانوار، بیش‌تر از ۴۴۴۶۶۶/۱ ریال و کمتر یا مساوی ۸۳۲۵۴۴/۹ ریال است ($Y = ۲$)	۳۶۹۱	۳۱/۸
هزینه‌ی ماهانه‌ی هر فرد خانوار، بیش‌تر از ۸۳۲۵۴۴/۹ ریال و کمتر یا مساوی ۱۳۳۸۴۳۴ ریال است ($Y = ۳$)	۲۹۰۶	۲۵
هزینه‌ی ماهانه‌ی هر فرد خانوار، بیش‌تر از ۱۳۳۸۴۳۴ ریال است ($Y = ۴$)	۲۹۰۴	۲۵

نتایج نشان می‌دهد که ۵۰ درصد افراد، فقیر یا شبه‌فقیر هستند. بیش‌ترین سهم مربوط به افراد شبه‌فقیر است و کمترین سهم مربوط به افراد فقیر. به‌منظور آزمون فرض برابری نسبت‌های چهار گروه مورد بررسی، $H_۰ : \pi_۱ = \pi_۲ = \pi_۳ = \pi_۴$، در مقابل این فرض که دست‌کم یکی از دو نسبت فوق مساوی نباشند، اگر از آماره‌ی خی‌دو به صورت

$$\chi^۲ = \sum_{i=۱}^{k} \frac{(O_i - E_i)^۲}{E_i}$$

استفاده کنیم که در آن E_i و O_i به‌ترتیب، مقدار مورد انتظار تحت $H_۰$ و مقدار مشاهده‌شده در هر طبقه‌ی فقر هستند و k تعداد خانه‌های مختلف جدول است ($k = ۴$ در جدول ۷.۱). برای داده‌های جدول ۷.۱، با توجه به آن‌که برای $i = ۱, ۲, ۳, ۴$ تحت $H_۰$ داریم $E_i = ۰/۲۵ \times ۱۱۶۱۹ = ۲۹۰۴/۷۵$، مقدار $\chi^۲_{\text{obs}}$ (خی‌دوِ مشاهده‌شده تحت جدول ۷.۱) برابر با ۴۲۵/۹۱۲ است. از آن‌جا که $\chi^۲$ تحت $H_۰$ دارای توزیع تقریبی خی‌دو با ۳ درجه‌ی آزادی است، P-مقدار برابر با ۰/۰۰۰ خواهد بود که از $P = \Pr_{H_۰}(\chi^۲_{k-۱} \geq \chi^۲_{\text{obs}})$-مقدار، محاسبه می‌شود و این کمّیت نشان می‌دهد که رده‌های پاسخ دارای نسبت‌های یکسان در جامعه نیستند. البته در این‌جا فقط برای تشریح چگونگی استفاده از $\chi^۲$ مثالی آوردیم. وقتی متغیر پاسخ، ترتیبی است، توصیه می‌شود که از $\chi^۲$ برای آزمون برابری نسبت‌ها استفاده نشود (اگرستی، ۲۰۰۲، ص. ۸۵).

به‌منظور بررسی الگوی وضعیت فقر در جامعه به تفکیک جنس، جدول ۸.۱ رده‌های مختلف وضعیت فقر را به تفکیک جنس نشان می‌دهد.

با توجه به اطلاعات موجود در این جدول می‌توان به بررسی توزیع فقر بین زنان و مردان

سرپرست خانوار پرداخت و این‌که آیا توزیع فقر سرپرستان زن و مرد، یکسان است یا نه. به همین دلیل از آماره‌ی d سامرز (سامرز، ۱۹۶۲) به‌منظور بررسی وجود پیوند بین یک متغیر دودویی و یک متغیر ترتیبی استفاده می‌کنیم.

جدول ۸.۱: جدول فراوانی و فراوانی نسبی سطری وضعیت فقر به تفکیک جنس در سال ۱۳۸۴

جنس	متغیر پاسخ (Y)				کل سطری
	۱	۲	۳	۴	
مرد	۱۹۳۵	۳۴۱۶	۲۶۷۶	۲۶۱۶	۱۰۶۴۳
فراوانی نسبی	(۰/۱۸۲)	(۰/۳۲۱)	(۰/۲۵۱)	(۰/۲۴۶)	(۱/۰۰)
زن	۱۸۳	۲۷۵	۲۳۰	۲۸۸	۹۷۶
فراوانی نسبی	(۰/۱۸۷)	(۰/۲۸۲)	(۰/۲۳۶)	(۰/۲۹۵)	(۱/۰۰)

این اندازه‌گیری، اختلاف بین جفت‌های هماهنگ و ناهماهنگ (در جفت‌های نابرابر) روی متغیر کمکی دودویی است (بخش ۲.۲.۲.۱ را ببینید). این آماره مقادیر [۱ ,۱−] را اخذ می‌کند (برای اطلاعات بیش‌تر به اگرستی، ۱۹۸۴، مراجعه شود). این کمّیت در نرم‌افزار SPSS از قسمت Analyze > Descriptive Statistics > Crosstabs به‌راحتی قابل محاسبه است. مقدار d سامرز برای جدول ۸.۱ برابر با ۰/۰۱۴− با انحراف معیار ۰/۰۰۷ (P−مقدار = ۰/۰۳۴) است که نشان می‌دهد بین جنس و پاسخ، همبستگی وجود دارد، یا به عبارت دیگر، توزیع فقر سرپرستان زن و مرد، متفاوت است. به‌منظور بررسی ارتباط بین جنس و وضعیت فقر در جدول ۸.۱ می‌توان از آماره‌های $\chi^۲$ و آزمون نسبت درست‌نمایی تعمیم‌یافته نیز استفاده کرد (مود و دیگران، ۱۹۷۴، فصل ۹). نتایج این آزمون بر اساس آماره‌ی خی‌دو مقدار ۱۴/۰۶۲ (d.f. = ۳، P−مقدار = ۰/۰۰۳) را نشان می‌دهد. همچنین بر اساس نتایج آزمون نسبت درست‌نمایی تعمیم‌یافته، مقدار این آماره برابر با $G^۲ = -۲[\log L_{\text{Null}} - \log L_{\text{Full}}] = ۱۳/۷۹۸$ است که در آن L_{Null} و L_{Full} توابع درست‌نمایی‌اند که به‌ترتیب در براورد ماکسیمم درست‌نمایی برای مدل تحت فرض صفر و مدل تحت فرض کلّی به دست آمده‌اند. $G^۲$ دارای توزیع خی‌دو (تحت این فرض که مدل تحت فرض صفر درست است) با ۳ درجه‌ی آزادی است. نتایج بررسی رابطه‌ی بین توزیع فقر و جنس سرپرستان خانوار با استفاده از سه روش معرفی‌شده نشان‌دهنده‌ی آن است که جنسیت، عامل مهمی در وضعیت فقر است.

۲.۲.۲.۱ بیکاری در طرح آمارگیری نیروی کار در سال ۱۳۸۵

در این مثال، از مجموعه‌داده‌ای استفاده شده که شامل افراد بیکار در فصل بهار ۱۳۸۵ است و از دادگان آمارگیری از نیروی کار مرکز آمار ایران استخراج شده است. در داده‌های اصلی، طول مدت بیکاری این افراد در بهار ۱۳۸۵ بر حسب ماه گردآوری شده است که با توجه به دلایل زیر، این پاسخ‌ها به‌صورت رده‌های ترتیبی طول مدت بیکاری بر حسب ماه در بازه‌های [۱,۳]، [۴,۶]، [۷,۱۲]، [۱۳,۲۴] و +۲۵، رده‌بندی شده‌اند. یکی از دلایل ترتیبی کردن طول مدت بیکاری این است که پاسخ‌های افراد معمولاً به‌طور تقریبی پاسخ‌های درستی هستند و ممکن است که این پاسخ‌ها همراه با خطا باشند. از طرفی برای سیاست‌گذاران و تحلیلگران بازار کار، توزیع زمان بیکاری در رده‌های طول مدت بیکاری برای ارائه‌ی تحلیل‌هایی در مورد بیکاری کوتاه‌مدت، میان‌مدت و بلندمدت حائز اهمیت است. جدول ۹.۱ توزیع نمونه‌ای طول مدت بیکاری بخشی از افراد بیکار در بهار ۱۳۸۵ را بیان می‌کند.

به‌منظور بررسی رابطه‌ی میزان تحصیلات و طول مدت بیکاری افراد بیکار با استفاده از اطلاعات این داده‌ها، جدول ۱۰.۱ طراحی شده است. نتایج جدول ۱۰.۱ نشان می‌دهد که در هر یک از رده‌های طول مدت بیکاری، عمده‌ی افراد، دارای تحصیلات زیر دیپلم هستند و کمترین سهم به افراد دارای تحصیلات عالی دانشگاهی برمی‌گردد. برای بررسی پیوند بین دو پاسخ ترتیبی می‌توان از کمّیت γ استفاده کرد که توسط گودمن و کروسکال (۱۹۵۴) پیشنهاد شده است. این کمّیت با نسبت اختلاف بین جفت‌های هماهنگ و ناهماهنگ بر مجموع جفت‌های هماهنگ (C) و ناهماهنگ (D) براورد و با نماد $\hat{\gamma} = \frac{C-D}{C+D}$ مشخص می‌شود. این مقدار، به «گامای نمونه‌ای» معروف است (با این نمادگذاری، d سامرز برابر است با $d = \dfrac{C-D}{\binom{n}{2} - T_X}$ که در آن n تعداد مشاهدات است و $T_X = \frac{\sum_{i=1}^{2} n_{i+}(n_{i+}-1)}{2}$) . مقدار گاما برای جامعه به‌صورت $\gamma = \frac{\Pi_c - \Pi_d}{\Pi_c + \Pi_d}$ تعریف می‌شود، که در آن Π_c و Π_d به‌ترتیب، احتمال‌های هماهنگی و ناهماهنگی برای یک جفت از مشاهدات انتخاب‌شده‌ی تصادفی‌اند. دامنه‌ی مقادیر ممکن گاما $[-1,1]$ است و زمانی مقدار ۱ را اخذ می‌کند که $\Pi_d = 0$ باشد. این آماره زمانی مقدار ۱− را اخذ می‌کند که $\Pi_c = 0$ باشد. این نکته قابل توجه است که $|\gamma| = 1$ دلالت بر

یکنوا بودن پیوند بین متغیرهای X و Y دارد، اما نه ضرورتاً اکیداً یکنوا (اگرستی، ۲۰۰۲، فصل ۲). برای اطلاعات جدول ۱۰.۱ مقدار آماره‌ی گاما برابر با ۰/۲۴۵ با انحراف معیار تقریبی ۰/۰۲۳ است (این مقدار انحراف معیار با استفاده از روش دلتا، δ، محاسبه می‌شود، که این روش در پیوست ۱ مورد بررسی قرار گرفته است) که نشان‌دهنده‌ی پیوند بالا بین متغیر پاسخ (طول مدت بیکاری) و سطح تحصیلات است. محاسبه‌ی این پیوند و پیوند بین دو پاسخ اسمی با استفاده از نرم‌افزارهای موجود، مانند SPSS از قسمت Crosstabs > Descriptive Statistics > Analyze و SAS با استفاده از Proc Freq امکان‌پذیر است.

جدول ۹.۱: توزیع طول مدت بیکاری

طول مدت بیکاری (بهار ۱۳۸۵)	فراوانی	درصد
[۱, ۳]	۵۳۰	۲۰/۹۹
[۴, ۶]	۴۱۸	۱۶/۵۵
[۷, ۱۲]	۵۷۵	۲۲/۷۷
[۱۳, ۲۴]	۴۸۹	۱۹/۳۷
۲۵+	۵۱۳	۲۰/۳۲
کل	۲۵۲۵	۱۰۰/۰

جدول ۱۰.۱: توزیع طول مدت بیکاری افراد بیکار و سطح تحصیلات

طول مدت بیکاری (بهار ۱۳۸۵)	سطح تحصیلات			
	زیر دیپلم	دیپلم و پیش‌دانشگاهی	تحصیلات عالی	
[۱, ۳]	۳۶۱ %۶۸/۱۱ %۲۹/۹۳	۱۱۳ %۲۱/۳۲ %۱۳/۸۸	۵۶ %۱۰/۵۷ %۱۱/۰۹	۵۳۰ %۱۰۰/۰۰ %۵۴/۹۰
[۴, ۶]	۲۲۷ %۵۴/۳۱ %۱۸/۸۲	۱۱۶ %۲۷/۷۵ %۱۴/۲۵	۷۵ %۱۷/۹۴ %۱۴/۸۵	۴۱۸ %۱۰۰/۰۰ %۴۷/۹۲
[۷, ۱۲]	۲۳۴ %۴۰/۷۰ %۱۹/۴۰	۲۰۷ %۳۶/۰۰ %۲۵/۴۳	۱۳۴ %۲۳/۳۰ %۲۶/۵۳	۵۷۵ %۱۰۰/۰۰ %۷۱/۳۶
[۱۳, ۲۴]	۱۸۸ %۳۸/۴۵ %۱۵/۵۹	۱۸۲ %۳۷/۲۲ %۲۲/۳۶	۱۱۹ %۲۴/۳۴ %۲۳/۵۶	۴۸۹ %۱۰۰/۰۰ %۶۱/۵۱
۲۵+	۱۹۶ %۳۸/۲۱ %۱۶/۲۵	۱۹۶ %۳۸/۲۱ %۲۴/۰۸	۱۲۱ %۲۳/۵۹ %۲۳/۹۶	۵۱۳ %۱۰۰/۰۰ %۶۴/۲۹
کل	۱۲۰۶ %۴۷/۷۶ %۱۰۰/۰۰	۸۱۴ %۳۲/۲۴ %۱۰۰/۰۰	۵۰۵ %۲۰/۰۰ %۱۰۰/۰۰	۲۵۲۵ %۱۰۰/۰۰ %۱۰۰/۰۰

۳.۲.۱ داده‌های مقطعی با پاسخ اسمی

در این زیربخش، یک کاربرد واقعی در مطالعات مقطعی برای پاسخ‌های اسمی با استفاده از طرح آمارگیری نیروی کار مرکز آمار ایران آورده می‌شود.

۱.۳.۲.۱ وضع فعالیت اقتصادی سرپرست خانوار در طرح آمارگیری نیروی کار در سال ۱۳۸۵

مثالی از مطالعات مقطعی اسمی در مطالعات خانوارهای ایرانی می‌تواند وضع فعالیت اقتصادی سرپرست خانوار با رده‌های بیکار، شاغل و غیر فعال باشد. جدول ۱۱.۱ که از نتایج طرح نیروی کار در سال ۱۳۸۵ به دست آمده، بخشی از داده‌های مربوط به تابستان ۱۳۸۵ است. در واقع، داده‌های جدول ۱۱.۱ وضع فعالیت اقتصادی افرادی در تابستان ۱۳۸۵ را نشان می‌دهد که در بهار ۱۳۸۵ بیکار بوده‌اند (اندازه‌ی نمونه = ۲۵۲۵).
در فصل ۴ و در هنگام معرفی مدل‌های مقطعی برای پاسخ‌های اسمی، از این داده‌ها استفاده خواهد شد.

جدول ۱۱.۱: وضع فعالیت افراد در فصل تابستان ۱۳۸۵ که در فصل بهار بیکار بوده‌اند

وضع فعالیت	فراوانی	درصد
بیکار	۹۶۰	۳۸/۰۲
شاغل	۸۸۲	۳۴/۹۳
غیر فعال	۶۸۳	۲۷/۰۵

۴.۲.۱ داده‌های طولی با پاسخ دودویی

در این زیربخش، سه کاربرد واقعی در مطالعات طولی با پاسخ دودویی آورده می‌شود.

۱.۴.۲.۱ داده‌های مربوط به آسم

مثالی از داده‌های طولی دودویی، داده‌های مربوط به آسم (رتنیتزکی و ویپیج، ۱۹۹۴) است. این داده‌ها در ۶ منطقه‌ی شهر هاروارد در ایالات اوهایوی آمریکا گردآوری شده‌اند، به این ترتیب که ۷۰۶ پسر و ۷۱۳ دختر سفیدپوست را در ۹ سالگی و دوباره در ۱۳ سالگی از نظر ابتلا یا عدم ابتلا به آسم، مورد مطالعه قرار داده‌اند. می‌خواهیم تأثیر

سن و جنس را بر ابتلا به آسم، مورد آزمون قرار دهیم.
جدول ۱۲.۱ داده‌های کامل و بدون مقادیر گم‌شده برای متغیر پاسخ را نشان می‌دهد. با استفاده از این داده‌ها در فصل‌های بعد نشان خواهیم داد که در نظر نگرفتن همبستگی پاسخ‌ها در سنین ۹ و ۱۳ سالگی چه تأثیری بر استنباط در مورد پارامترهای رگرسیونی می‌گذارد. همان‌طور که جدول ۱۲.۱ نشان می‌دهد، از ۷۰۶ پسر، ۵۵۷ نفر و از ۷۱۳ دختر، ۵۹۰ نفر به متغیر مورد علاقه پاسخ داده‌اند که از میان پسرها، ۲۲ نفر در هر دو سن ۹ و ۱۳ سالگی آسم داشته‌اند در حالی که از میان دخترها، ۱۳ نفر در هر دو سن ۹ و ۱۳ سالگی آسم داشته‌اند.

جدول ۱۲.۱: داده‌های مربوط به آسم (منبع: رتنیتزکی و ویپیج، ۱۹۹۴)

پسرها		۱۳سالگی		جمع کل
		آسم دارد	آسم ندارد	
۹ سالگی	آسم دارد	۲۲	۶	۲۸
	آسم ندارد	۱۵	۵۱۴	۵۲۹
	جمع کل	۳۷	۵۲۰	۵۵۷
		۱۳سالگی		
دخترها		آسم دارد	آسم ندارد	جمع کل
۹ سالگی	آسم دارد	۱۳	۳	۱۶
	آسم ندارد	۱۳	۵۶۱	۵۷۴
	جمع کل	۲۶	۵۶۴	۵۹۰

۲.۴.۲.۱ اثر آلودگی هوا بر سلامت کودکان

مثال دیگری از داده‌های طولی دودویی، مثال اثر آلودگی هوا بر سلامت کودکان است (ویر و دیگران، ۱۹۸۴) که در شش شهر آلوده‌ی آمریکا گردآوری شده است (جدول ۱۳.۱). در این مثال، ۵۳۷ کودک را در این شش شهر، مورد بررسی قرار داده و بیمار بودن یا نبودن آنان را در سنین ۷ تا ۱۰ سالگی مشاهده کرده‌اند (۱: بیمار بودن و ۰: بیمار نبودن). در این مثال، اثر سیگار کشیدن مادر بر روی سلامت کودک‌اش مورد بررسی قرار گرفته است. به‌عنوان مثال، تعداد کودکانی که در سنین ۷، ۸، ۹ و ۱۰ سالگی بیمار نبوده‌اند و مادرشان نیز سیگاری نبوده، ۲۳۷ نفر گزارش شده است. به همین ترتیب می‌توان بقیه‌ی اعداد داخل خانه‌های جدول را تعبیر کرد.

جدول ۱.۱۳: داده‌های شش شهر برای بیماری کودکان (بیماری: بلی، عدم بیماری: نه)

			۱۰ سالگی			
			مادر غیرسیگاری		مادر سیگاری	
۷ سالگی	۸ سالگی	۹ سالگی	نه	بلی	نه	بلی
نه	نه	نه	۲۳۷	۱۰	۱۱۸	۶
		بلی	۱۵	۴	۸	۲
	بلی	نه	۱۶	۲	۱۱	۱
		بلی	۷	۳	۶	۴
بلی	نه	نه	۲۴	۳	۷	۳
		بلی	۳	۲	۳	۱
	بلی	نه	۶	۲	۴	۲
		بلی	۵	۱۱	۴	۷

۳.۴.۲.۱ داده‌های بی‌خوابی

مثال دیگری از پاسخ‌های دودویی طولی، داده‌های بی‌خوابی است. بیماری بی‌خوابی یک اختلال و آشفتگی در خواب است، به‌گونه‌ای که بیمار به اندازه‌ی کافی نمی‌خوابد و از خوابیدن رضایت ندارد. بی‌خوابی می‌تواند با توجه به مدت زمانی که طول می‌کشد و این‌که چگونه اتفاق می‌افتد، متفاوت باشد. بی‌خوابی می‌تواند کوتاه‌مدت (بی‌خوابی حاد) یا بلندمدت (بی‌خوابی مزمن) باشد. بی‌خوابی حاد می‌تواند از یک شب تا چند هفته طول بکشد. بی‌خوابی مزمن زمانی اتفاق می‌افتد که یک فرد دست‌کم در ۳ شب از یک هفته برای مدت یک ماه یا طولانی‌تر، بی‌خوابی داشته باشد، که دلیل آن می‌تواند عوامل متعددی باشد و اغلب با مشکلات دیگری نیز همراه است. دلایل عمده‌ی بی‌خوابی مزمن، افسردگی، استرس‌های زیاد و درد یا ناراحتی در شب است.

بی‌خوابی حاد ممکن است نیاز به درمان نداشته باشد. درمان بی‌خوابی مزمن ابتدا از طریق درمان شرایط اساسی یا مشکلات سلامت که منجر به بی‌خوابی می‌شود، صورت می‌گیرد. اگر بی‌خوابی ادامه پیدا کند، پزشک، درمان‌های رفتاری یا درمان دارویی را پیشنهاد می‌کند. در این مثال، اندازه‌ی مورد نظر پاسخ بیمار به سؤال «با چه سرعتی بعد از این‌که به رختخواب می‌روید می‌خوابید؟» به‌صورت کمتر از ۲۰ دقیقه، ۳۰—۲۰ دقیقه، ۶۰—۳۰ دقیقه و بیش‌تر از ۶۰ دقیقه طبقه‌بندی شده است. این سؤال از بیماران بعد از یک دوره‌ی درمانی یک‌هفته‌ای که از دارونما (اندازه‌گیری مبنا) استفاده کرده‌اند و به دنبال آن، بعد از یک دوره‌ی درمانی دو هفته‌ای پرسیده شده است (فرانکوم و دیگران،

۱۹۸۹). داده‌هایی که در فصل ۲ از آن استفاده خواهیم کرد، داده‌های بی‌خوابی است، که متغیر پاسخ ترتیبی آن را به متغیر دودویی کمتر از ۳۰ دقیقه و بیش‌تر از ۳۰ دقیقه رده‌بندی کرده‌ایم. دلیل ادغام بعضی از سطوح پاسخ‌های ترتیبی اولیه این است که رده‌بندی پاسخ ترتیبی زمان دوم به‌وسیله‌ی نوع درمان و پاسخ ترتیبی زمان اول منجر به داشتن برخی خانه‌های دارای مقدار صفر می‌شود و بعضی از خانه‌ها نیز کمتر از ۵ مشاهده خواهند داشت. با توجه به ادغام سطوح پاسخ‌ها، تحلیل‌های توصیفی روی داده‌های بی‌خوابی، کمی تغییر می‌کند (جدول ۱۴.۱ را ببینید).
جدول ۱۵.۱ توزیع‌های حاشیه‌ای نمونه‌ای را برای پاسخ‌های زمان اول و دوم در دو نوع درمان نشان می‌دهد. احتمال‌های حاشیه‌ای نمونه‌ای در جدول ۱۵.۱ نشان می‌دهند که در زمان اول، نسبت بیمارانی که از داروی مؤثر استفاده می‌کنند و به‌کندی به خواب می‌روند ($\geq$ ۳۰)، برابر با ۰/۷۳۱ است و این نسبت بعد از مصرف دارو به ۰/۲۵۲ کاهش می‌یابد. در واقع، این بدان معنا است که با استفاده از این دارو، بیماران با سرعت بیش‌تری به خواب می‌روند.

جدول ۱۴.۱: مدت زمان قبل از به خواب رفتن (بر حسب دقیقه) که از پرسیدن سؤال «به چه سرعت خواب‌تان می‌برد؟» به دست می‌آید (پاسخ زمان دوم، بر اساس نوع درمان و پاسخ زمان اول، تعداد و درصد سطرها. منبع: فرانکوم و دیگران، ۱۹۸۹).

نوع درمان	پاسخ در زمان اول (Y_1)	پاسخ در زمان دوم (Y_2)		مجموع
		$<$ ۳۰	$\geq$ ۳۰	
مؤثر	$<$ ۳۰	۲۷	۵	۳۲
	درصد	۸۴/۴	۱۵/۶	۱۰۰/۰
	$\geq$ ۳۰	۶۲	۲۵	۸۷
	درصد	۷۱/۳	۲۸/۷	۱۰۰/۰
دارونما	$<$ ۳۰	۳۰	۴	۳۴
	درصد	۸۸/۲	۱۱/۸	۱۰۰/۰
	$\geq$ ۳۰	۳۰	۵۶	۸۶
	درصد	۳۴/۹	۶۵/۱	۱۰۰/۰

جدول ۱۵.۱: توزیع حاشیه‌ای نمونه‌ای برای پاسخ‌های اولیه و ثانویه برای دو روش درمان

پاسخ	نوع درمان	طبقه‌بندی پاسخ	
		$<$ ۳۰	$\geq$ ۳۰
پاسخ اولیه	مؤثر	۰/۲۶۹	۰/۷۳۱
	دارونما	۰/۲۸۴	۰/۷۱۶
پاسخ ثانویه	مؤثر	۰/۷۴۸	۰/۲۵۲
	دارونما	۰/۵۰۰	۰/۵۰۰

مطالعات زیادی برای مقایسه‌ی گروه‌ها روی متغیرهای پاسخ دودویی طراحی شده است. به‌عنوان مثال، اگر متغیر پاسخ، دارای دو رده‌ی موفقیت و شکست برای پاسخ به

یک نوع درمان دارویی باشد، با در نظر گرفتن دو گروه از افراد، یک جدول پیشایندی ۲×۲ داریم که در آن، سطرها می‌توانند گروه‌ها باشند و ستون‌ها رده‌های پاسخ. برای مقایسه‌ی گروه‌های مختلف می‌توانیم در این جدول پیشایندی ۲×۲، مقایسه‌هایی انجام دهیم. به‌عنوان مثال، در مثال بی‌خوابی، اثر دو نوع درمان را روی متغیر پاسخ (مدت زمان به خواب رفتن)، در نظر می‌گیریم. بنا بر این مثلاً برای پاسخ در زمان اول، سطرها دو نوع روش درمانی و ستون‌ها رده‌های پاسخ در زمان اول هستند (جدول ۱۶.۱ را ببینند). اکنون مقایسه‌هایی برای گروه‌های مختلف در این مثال انجام می‌دهیم.

اختلاف نسبت‌ها $(\pi_{۱|i} - \pi_{۱|i'})$ برای هر i و i' یک مقایسه‌ی ساده برای هر دو رده‌ی مطلوب از یک جدول پیشایندی $۲ \times I$ می‌باشد که در مثال مورد نظر، سطرها دو روش درمانی‌اند $(I = ۲)$. اختلاف نسبت‌ها اعداد بین ۱− و ۱ را اخذ می‌کند. برای $I = ۲$ زمانی این اختلاف، مقدار صفر را اخذ می‌کند که متغیر پاسخ به‌طور آماری مستقل از طبقه‌بندی سطرها در مثال روش‌های درمانی باشد.

جدول ۱۶.۱: توزیع حاشیه‌ای دو روش درمان بر حسب رده‌های پاسخ در زمان اول

نوع درمان	طبقه‌بندی پاسخ		مجموع
	< ۳۰	≥ ۳۰	
مؤثر	۳۲	۸۷	۱۱۹
فراوانی نسبی	۰/۲۶۹	۰/۷۳۱	۰/۱۰۰
دارونما	۳۴	۸۶	۱۲۰
فراوانی نسبی	۰/۲۸۳	۰/۷۱۷	۰/۱۰۰

در این مثال، براورد اختلاف نسبت دو روش درمانی برای افرادی که طول مدت بی‌خوابی آن‌ها در زمان اول، کمتر از ۳۰ دقیقه است، برابر است با $-۰/۰۱۴ = ۰/۲۶۹ - ۰/۲۸۳ = \frac{۳۲}{۱۱۹} - \frac{۳۴}{۱۲۰}$، که این اختلاف، نزدیک صفر است و نشان می‌دهد پاسخ افراد در زمان اول ممکن است مستقل از نوع درمان باشد (البته برای آزمون‌های معناداری آماری باید واریانس براوردگر محاسبه شود). مقایسه‌ی پاسخ‌های افرادی که مدت به خواب رفتن آن‌ها بیش‌تر از ۳۰ دقیقه است نیز مشابه مقایسه‌ی قبل برای پاسخ‌های افرادی است که مدت بی‌خوابی آن‌ها کمتر از ۳۰ دقیقه می‌باشد؛ زیرا $۰/۰۱۴ = \hat{\pi}_{۱|۱} - \hat{\pi}_{۱|۲} = (۱ - \hat{\pi}_{۱|۲}) - (۱ - \hat{\pi}_{۱|۱}) = \hat{\pi}_{۲|۲} - \hat{\pi}_{۲|۱}$.

در مطالعاتی که مقایسه‌ی دو روش درمانی را روی نسبت آزمودنی‌هایی که نادرند مد نظر قرار می‌دهیم، اختلاف بین مثلاً ۰/۰۱۰ و ۰/۰۰۱ ممکن است خیلی مهم‌تر از اختلاف

بین ۰/۴۱۰ و ۰/۴۰۱ باشد اگرچه این اختلاف برای هر دو ۰/۰۰۹ است. در چنین مطالعاتی مخاطره‌ی نسبی، آگاهی بیش‌تری می‌دهد. مخاطره‌ی نسبی برای دو رده‌ی i و i' که به‌صورت $\frac{\pi_{1|i}}{\pi_{1|i'}}$ تعریف می‌شود، می‌تواند هر مقدار حقیقی نامنفی را اخذ کند. مخاطره‌ی نسبی ۱ به معنای آن است که هر دو روش درمان، تأثیری یکسان دارند. در مثالی که بیان کردیم، براورد مخاطره‌ی نسبی به‌ترتیب، $10 = \frac{0/010}{0/001}$ و $1/02 = \frac{0/410}{0/401}$ است که نشان‌دهنده‌ی به دست آوردن آگاهی بیش‌تر نسبت به استفاده از اختلاف نسبت‌ها می‌باشد (این وقتی اتفاق می‌افتد که $\pi_{1|i}$ و $\pi_{1|i'}$، هر دو نزدیک به صفر یا ۱ هستند). مقایسه‌ی روش‌های مختلف درمان با استفاده از رده‌ی دوم پاسخ، مخاطره‌ی نسبی متفاوتی (بر خلاف اختلاف نسبت‌ها) از مخاطره‌ی نسبی رده‌ی اول پاسخ می‌دهد، چرا که $\frac{\pi_{1|1}}{\pi_{1|2}} \neq \frac{1-\pi_{1|1}}{1-\pi_{1|2}}$.

همان‌طور که قبلاً نیز بیان شد، برای احتمال موفقیت π، بخت به‌صورت $\frac{\pi}{1-\pi}$ تعریف می‌شود که یک مقدار نامنفی است. در یک جدول پیشایندی 2×2 با سطر i برای $i = 1, 2$، بخت موفقیت به شکست به‌صورت $\frac{\pi_{1|i}}{1-\pi_{1|i}}$ نمایش داده می‌شود. نسبت بخت‌ها در دو سطر iام و i'ام که یکی دیگر از معیارهای مقایسه‌ی دو نسبت است، به‌صورت

$$\theta = \frac{\frac{\pi_{1|i}}{1-\pi_{1|i}}}{\frac{\pi_{1|i'}}{1-\pi_{1|i'}}}$$

تعریف می‌شود، که به آن نسبت بخت‌ها گفته می‌شود.

برای توزیع‌های توأم با احتمال‌های خانه‌ای π_{ij} برای $i, j = 1, 2$، تعریف معادل بخت در سطر iام به‌صورت $\frac{\pi_{1|i}}{\pi_{2|i}}$ برای $i = 1, 2$ است (جدول ۱۷.۱ را ببینید). بنا بر این نسبت بخت به‌صورت

$$\theta = \frac{\frac{\pi_{1|1}}{\pi_{2|1}}}{\frac{\pi_{1|2}}{\pi_{2|2}}} = \frac{\pi_{1|1}\pi_{2|2}}{\pi_{1|2}\pi_{2|1}} = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}$$

تعریف می‌شود. در مثال بی‌خوابی، براورد بخت افرادی که در زمان اول از روش درمان داروی مؤثر استفاده می‌کنند، برابر با $0/368 = \frac{0/269}{0/731}$ است. نسبت بخت براورد شده برای دو نوع روش درمان در زمان اول، برابر با $0/932 = \frac{0/269 \times 0/717}{0/731 \times 0/283}$ است. یک نام دیگر برای نسبت بخت‌ها، نسبت حاصل‌ضرب متقاطع است؛ زیرا این نسبت، معادل نسبت

حاصل‌ضرب $\pi_{11}\pi_{22}$ و $\pi_{12}\pi_{21}$ در جدول ۱۷.۱ است که از حاصل‌ضرب احتمال‌های سلول‌هایی به دست آمده‌اند که به‌صورت قطری در خلاف راستای یکدیگر هستند.

جدول ۱۷.۱: نمادگذاری برای احتمال‌های توأم و حاشیه‌ای

سطر	ستون ۱	ستون ۲	مجموع
۱	π_{11}	π_{12}	π_{1+}
۲	π_{21}	π_{22}	π_{2+}
مجموع	π_{+1}	π_{+2}	

استقلال دو متغیر X و Y (که برای سطرها و ستون‌ها تعریف شده بودند) به $\theta = 1$ می‌انجامد. برای $1 < \theta < \infty$ افرادی که در سطر اول هستند بخت موفقیت بیش‌تری دارند تا افرادی که در سطر دوم هستند. به‌عنوان مثال، $\theta = 4$ به این معنا است که بخت موفقیت در سطر اول، چهار برابر بخت موفقیت در سطر دوم است. این مطلب به آن معنا نیست که $\pi_{1|1} = 4\pi_{1|2}$ باشد، چرا که این به معنای مخاطره‌ی نسبی ۴ است. زمانی که $0 < \theta < 1$ است، بخت موفقیت در سطر دوم، بیش‌تر از بخت موفقیت در سطر اول است. زمانی که یکی از سلول‌ها احتمال صفر را اخذ کند، θ صفر یا ∞ می‌شود.

مقادیر θی دورتر از یک (در هر دو جهت بیش‌تر یا کمتر)، نشان‌دهنده‌ی همبستگی بیش‌تر است. وقتی $\theta = 0/25$ است، بخت موفقیت در سطر اول، $0/25$ برابر بخت موفقیت در سطر دوم است، یا به‌طور معادل، بخت موفقیت در سطر دوم، ۴ $(= \frac{1}{0/25})$ برابر بخت موفقیت در سطر اول است. نسبت بخت‌ها زمانی که جهت جدول بر عکس شود به‌گونه‌ای که جای سطرها و ستون‌ها عوض شود، تغییر نمی‌کند. در واقع، لزومی ندارد که هر طبقه‌بندی که متغیرها را به‌عنوان پاسخ یا متغیر تبیینی در نظر می‌گیرد، مد نظر قرار گیرد. به عبارت دیگر،

$$\theta = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} = \frac{\frac{\Pr(Y=1|X=1)}{\Pr(Y=2|X=1)}}{\frac{\Pr(Y=1|X=2)}{\Pr(Y=2|X=2)}} = \frac{\frac{\Pr(X=1|Y=1)}{\Pr(X=2|Y=1)}}{\frac{\Pr(X=1|Y=2)}{\Pr(X=2|Y=2)}}.$$

در حقیقت، نسبت بخت‌ها می‌تواند برای طرح‌های نمونه‌گیری گذشته‌نگر، آینده‌نگر و مقطعی به‌عنوان معیاری برای همبستگی دو متغیر، مورد استفاده قرار گیرد.

برای شمارش‌های خانه‌ای $\{n_{ij}\}$ به‌ازای $i, j = 1, 2$، نسبت بخت‌های نمونه‌ای، $\hat{\theta} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$ است. ارتباط بین نسبت بخت‌ها و مخاطره‌ی نسبی به‌صورت زیر است:

$$\text{نسبت بخت‌ها} = \text{مخاطره‌ی نسبی} \times \frac{۱-\pi_{۱|۲}}{۱-\pi_{۱|۱}}.$$

اکنون بازه‌ی اطمینان برای θ با استفاده از براوردگر نسبت بخت‌ها را برای یک جدول پیشایندی ۲ × ۲ به دست می‌آوریم. درستی براوردگرهای پارامترهای همبستگی به‌وسیله‌ی انحراف معیار توزیع نمونه‌گیری آن‌ها مشخص می‌شود. بنا بر این به‌منظور ارائه‌ی بازه‌ی اطمینان برای نسبت بخت‌ها، لازم است انحراف معیار نسبت بخت‌های نمونه‌ای را برای نمونه‌های بزرگ معرفی کنیم.

براورد نسبت بخت‌ها، $\hat{\theta} = \frac{n_{۱۱}n_{۲۲}}{n_{۱۲}n_{۲۱}}$، برای جدول پیشایندی ۲ × ۲ زمانی که $n_{۱۱} = ۰$ باشد، یا $n_{۲۲} = ۰$، معادل ۰ است و زمانی که $n_{۱۲} = ۰$ باشد، یا $n_{۲۱} = ۰$، برابر با ∞ است. چون این پاسخ‌ها دارای احتمال‌های مثبت هستند، مقادیر مورد انتظار و واریانس $\hat{\theta}$ و $\log\hat{\theta}$ وجود ندارد. با در نظر گرفتن معیارهای اریبی و میانگین توان دوم، گارت و تسوآیفون (۱۹۶۷) و هالدن (۱۹۵۶) نشان دادند که براوردگر اصلاح‌شده‌ی

$$\tilde{\theta} = \frac{(n_{۱۱} + ۰/۵)(n_{۲۲} + ۰/۵)}{(n_{۱۲} + ۰/۵)(n_{۲۱} + ۰/۵)}$$

برای θ، و براوردگر $\log\tilde{\theta}$ برای $\log\theta$ رفتار خوبی دارد. براوردگرهای $\hat{\theta}$ و $\tilde{\theta}$ توزیع نرمال مجانبی مشابهی حول θ دارند. به‌جز زمانی که n خیلی کوچک است، توزیع آن‌ها متقارن است و تبدیل لگاریتمی با سرعت بیش‌تری همگرا به توزیع نرمال است. روش δ (اگرستی، فصل ۳، ص. ۷۳؛ به پیوست ۱ نیز مراجعه شود) یک انحراف معیار مجانبی براورد شده برای $\log\hat{\theta}$ به‌صورت

$$\hat{\sigma}(\log\hat{\theta}) = \left(\frac{۱}{n_{۱۱}} + \frac{۱}{n_{۱۲}} + \frac{۱}{n_{۲۱}} + \frac{۱}{n_{۲۲}}\right)^{\frac{۱}{۲}}$$

فراهم می‌کند و چون $\log\hat{\theta}$ به‌طور مجانبی دارای توزیع نرمال است،

$$\log\hat{\theta} \pm z_{\frac{\alpha}{۲}}\hat{\sigma}(\log\hat{\theta})$$

یک بازه‌ی اطمینان مجانبی $(۱-\alpha)$ ۱۰۰ درصدی برای $\log\theta$ است که در آن، $z_{\frac{\alpha}{۲}}$ نقطه‌ای از توزیع نرمال استاندارد است که احتمال دُم سمت راست آن برابر با $\frac{\alpha}{۲}$ است. با گرفتن آنتی‌لگاریتم از نقاط انتهایی بازه‌ی اطمینان فوق، یک بازه‌ی اطمینان برای θ به دست می‌آید. ولف (۱۹۵۵) این بازه‌ی اطمینان را پیشنهاد کرد. زمانی که $\hat{\theta} = ۰$ یا

$\hat{\theta} = \infty$ باشد، بازه‌ی ولف وجود ندارد. زمانی که $\hat{\theta} = ۰$ باشد می‌توان ۰ را به‌عنوان حد پایین در نظر گرفت و زمانی که $\hat{\theta} = \infty$ باشد می‌توان ∞ را به‌عنوان حد بالا در نظر گرفت. حد دیگر را می‌توان با استفاده از فرمول ولف با إعمال برخی اصلاحات مانند روش گارت (۱۹۶۶) که $\{n_{ij}\}$ را با $\{n_{ij} + ۰/۵\}$ در براوردگر و انحراف معیار آن جایگزین می‌کند، به دست آورد. روش‌های دیگر برای یافتن بازه‌ی اطمینان، همچون معکوس کردن آزمون نمره (کورن فیلد، ۱۹۵۶) یا آزمون نسبت درست‌نمایی نیز می‌توانند برای θ به کار روند.

همان‌طور که قبلاً اشاره شد، اختلاف نسبت‌ها و مخاطره‌ی نسبی در یک جدول پیشایندی ۲×۲، توزیع شرطی یک متغیر پاسخ را برای دو گروه مقایسه می‌کنند. برای این اندازه‌ها، نمونه‌ها به‌عنوان دوجمله‌ای‌های مستقل در نظر گرفته می‌شوند و برای گروه iام، Y_i دارای توزیع دوجمله‌ای با اندازه‌ی نمونه‌ی n_i و احتمال $\pi_{۱|i}$ برای پاسخ موفقیت در نظر گرفته می‌شود. نسبت نمونه‌ای $\hat{\pi}_{۱|i} = \frac{y_{i۱}}{n_{i+}}$ دارای امید ریاضی $\pi_{۱|i}$ و واریانس $\frac{\pi_{۱|i}(۱-\pi_{۱|i})}{n_{i+}}$ است که در آن برای $i = ۱, ۲$ داریم $n_{i+} = n_{i۱} + n_{i۲}$. فرض کنید $\hat{\pi}_{۱|۱}$ و $\hat{\pi}_{۱|۲}$ مستقل‌اند و اختلاف آن‌ها دارای میانگین و انحراف معیار زیر است:

$$E(\hat{\pi}_{۱|۱} - \hat{\pi}_{۱|۲}) = \pi_{۱|۱} - \pi_{۱|۲}$$
$$\sigma(\hat{\pi}_{۱|۱} - \hat{\pi}_{۱|۲}) = \left[\frac{\pi_{۱|۱}(۱-\pi_{۱|۱})}{n_{۱+}} + \frac{\pi_{۱|۲}(۱-\pi_{۱|۲})}{n_{۲+}}\right]^{\frac{۱}{۲}}$$

براوردگر $\hat{\sigma}(\hat{\pi}_{۱|۱} - \hat{\pi}_{۱|۲})$ با جایگذاری $\hat{\pi}_{۱|i}$ به‌جای $\pi_{۱|i}$ در رابطه‌ی فوق به دست می‌آید. بنا بر این بازه‌ی اطمینان

$$(\hat{\pi}_{۱|۱} - \hat{\pi}_{۱|۲}) \pm z_{\frac{\alpha}{۲}}\hat{\sigma}(\hat{\pi}_{۱|۱} - \hat{\pi}_{۱|۲})$$

یک بازه‌ی اطمینان $(۱-\alpha)$ ۱۰۰ درصدی برای $\pi_{۱|۱} - \pi_{۱|۲}$ است.

لگاریتم مخاطره‌ی نسبی نمونه‌ای، که با $\log(r) = \log(\frac{\hat{\pi}_{۱|۱}}{\hat{\pi}_{۱|۲}})$ نمایش داده می‌شود، مشابه نسبت بخت‌ها سریع‌تر از خود مخاطره‌ی نسبی به نرمال بودن همگرا است. انحراف معیار $\log(r)$ با استفاده از روش δ به‌صورت زیر است:

$$\sigma(\log(r)) = \left(\frac{۱-\pi_{۱|۱}}{\pi_{۱|۱} \times n_{۱+}} + \frac{۱-\pi_{۱|۲}}{\pi_{۱|۲} \times n_{۲+}}\right)^{\frac{۱}{۲}}.$$

بنا بر این بازه‌ی اطمینان برای لگاریتم مخاطره‌ی نسبی به‌صورت زیر است:

$$\log(r) \pm z_{\frac{\alpha}{۲}}\hat{\sigma}(\log(r)),$$

که با گرفتن آنتی‌لگاریتم از نقاط انتهایی رابطه‌ی فوق، بازه‌ی اطمینان برای مخاطره‌ی نسبی به دست می‌آید.

بر اساس داده‌های بی‌خوابی در جدول ۱۶.۱، برای افرادی که از داروی مؤثر استفاده می‌کنند، نسبت نمونه‌ایِ به خواب رفتن در کمتر از ۳۰ دقیقه ۰/۲۶۹ است و نسبت مشابه برای افرادی که از دارونما استفاده می‌کنند ۰/۲۸۳ است. مخاطره‌ی نسبی نمونه‌ای برابر با $۰/۹۵۱ = \frac{۰/۲۶۹}{۰/۲۸۳}$ است و بازه‌ی اطمینان ۹۵ درصدی برای لگاریتم مخاطره‌ی نسبی، برابر است با $\log(۰/۹۵۱) \pm ۱/۹۶(\frac{۰/۷۳۱}{۰/۲۶۹\times ۱۱۹} + \frac{۰/۷۱۷}{۰/۲۸۳\times ۱۲۰}) = (-۰/۱۳۶, ۰/۰۳۶)$. این بازه‌ی اطمینان برای مخاطره‌ی نسبی به $(۰/۸۹۷, ۱/۰۶۲)$ تبدیل می‌شود. بنا بر این می‌توان نتیجه گرفت که میزان بی‌خوابی کمتر از ۳۰ دقیقه برای افرادی که از داروی مؤثر استفاده می‌کنند، بین ۰/۸۹۷ و ۱/۰۶۲ برابر افرادی است که از دارونما استفاده می‌کنند.

بازه‌ی اطمینان ۹۵ درصدی برای اختلاف $\pi_{1|2} - \pi_{1|1}$ برابر است با

$$(۰/۲۸۳ - ۰/۲۶۹) \pm ۱/۹۶\left(\frac{۰/۷۳۱ \times ۰/۲۶۹}{۱۱۹} + \frac{۰/۷۱۷ \times ۰/۲۸۳}{۱۲۰}\right)^{\frac{۱}{۲}}$$

یا $(-۰/۰۹۹, ۰/۱۲۷)$. همچنین بخت نمونه‌ای در مثال بی‌خوابی برابر است با $\hat{\theta} = \frac{۳۲\times ۸۶}{۸۷\times ۳۴} = ۰/۹۳۰$. برای $\log\theta$ داریم $\log\hat{\theta} \pm z_{\frac{\alpha}{۲}}\hat{\sigma}(\log\hat{\theta}) = (-۰/۶۳۹, ۰/۴۹۵)$ که به بازه‌ی اطمینان ۹۵ درصدی $(۰/۵۲۸, ۱/۶۴۱)$ برای نسبت بخت‌ها تبدیل می‌شود. تمامی معیارهای پیوند و بازه‌های اطمینان مربوط، گواهی می‌دهند که در زمان اول، تفاوتی در استفاده از داروی مؤثر یا دارونما وجود ندارد.

در مثال داده‌های بی‌خوابی، نسبت بخت براورد شده برای دو پاسخ (بدون توجه به تأثیر درمان) در زمان اول و دوم، ۵/۵۷۶ است. بنا بر این همبستگی زیادی بین دو پاسخ وجود دارد و در تحلیل این داده‌ها باید این همبستگی در نظر گرفته شود. مقدار آماره‌ی d سامرز برای بررسی پیوند بین دو پاسخ، ۰/۳۰۵ است که نشان‌دهنده‌ی همبستگی زیاد بین دو پاسخ است. همچنین نسبت بخت براورد شده بین دو پاسخ، برای افرادی که از داروی مؤثر و دارونما استفاده می‌کنند، به‌ترتیب برابر با ۲/۱۷ (با انحراف معیار ۰/۱۳۴) و ۱۴ (با انحراف معیار ۰/۴۷۸) است. این اعداد نشان می‌دهند که پیوند بین پاسخ‌ها با سطوح مختلف درمان تغییر می‌کند. بنا بر این در نظر گرفتن اثر درمان در ساختار همبستگی، امری

ضروری است.

۵.۲.۱ داده‌های طولی با پاسخ ترتیبی

در این زیربخش، دو کاربرد واقعی در مطالعات طولی با پاسخ ترتیبی آورده می‌شود.

۱.۵.۲.۱ بررسی اثر جانبی استفاده از داروی فلوکسامین

یک مثال کاربردی برای پاسخ‌های ترتیبی طولی می‌تواند بررسی اثر جانبی استفاده از داروی فلوکسامین باشد. از این داده‌ها در یک مطالعه‌ی چندمرکزی گردآوری شده‌اند که در آن، داروی فلوکسامین را ۳۱۵ فرد مصرف کرده‌اند. این داده‌ها در مولنبرگ و لسافره (۱۹۹۴)، کن‌وارد و دیگران (۱۹۹۴) و مولنبرگ و دیگران (۱۹۹۷) استفاده شده است. این مثال، یک مطالعه‌ی چهار دوره‌ای است که در اولین مراجعه، اطلاعات متغیرهای کمکی پایه، مانند جنس، سن، شدت بیماری (مقیاس ۱ تا ۷) و طول مدت بیماری ذهنی برای هر آزمودنی اندازه‌گیری شده است. بعد از مراجعه، آزمودنی‌ها ۴ معاینه در هفته‌های ۲ و ۴ و ۸ و ۱۲ داشتند. در هر معاینه، آثار جانبی پاسخ‌های ترتیبی با کدهای زیر ثبت شده است:

۰: بدون اثر جانبی، ۱: بدون اختلال معنادار با عملکرد بیمار، ۲: اختلال معنادار با عملکرد بیمار، و ۳: اثر جانبی مقدم بر اثر درمانی.

از آن‌جا که برخی از بیماران در این مطالعه دارای پاسخ گم‌شده به‌صورت انصراف هستند (به این معنا که در زمانی خاص، از مطالعه خارج شده‌اند و دیگر به آن برنگشته‌اند)، در این کتاب و در فصل ۷، مشابه هینس و هینس (۲۰۰۵) و برای اطمینان از این‌که فرض انصراف تصادفی برقرار باشد، از داده‌ها در سه دوره‌ی زمانی هفته‌های ۲، ۴ و ۱۲ استفاده خواهیم کرد. از افرادی که برای اولین بار مراجعه کرده‌اند، ۱۴ نفر اصلاً مراجعه‌ی دیگری نداشته‌اند، ۲ نفر بی‌پاسخی غیر یکنوا داشته‌اند (به این معنا که در زمانی از مطالعه خارج شده‌اند و دوباره در زمانی دیگر به آن بازگشته‌اند) و ۱۲ نفر دست‌کم در یکی از متغیرهای کمکی سن، مدت بیماری و شدت اولیه‌ی بیماری ذهنی، پاسخ گم‌شده داشته‌اند. بنا بر این با حذف این موارد، ۲۸۷ نفر در تحلیل باقی خواهند ماند. از طرفی چون رده‌ی ۳

پاسخ، فقط در ۲ درصد از مواقع رخ داده بود، برای سادگی محاسبات، این رده با رده‌ی ۲ پاسخ ادغام خواهد شد. تعداد الگوهای مختلف داده‌های گم‌شده برای این داده‌ها به‌صورت OOO= ۲۱٦، OOM= ۴۳ و OMM= ۲۸ است که در آن O و M به‌ترتیب، معرف داده‌های مشاهده‌شده و گم‌شده‌اند. جدول ۱.۱۸ احتمال‌های نمونه‌ای هر سطح $Y_۱$ و $Y_۲$ و $Y_۳$ را به‌ترتیب برای ۲۸۷ و ۲۵۹ و ۲۱٦ بیمار نشان می‌دهد که در زمان‌های اول و دوم و سوم مشاهده شده‌اند. این جدول نشان می‌دهد که نسبت بیمارانی که در سطح ۲ مشاهده شده‌اند، کمتر از سطوح دیگر برای همه‌ی دوره‌ها است. احتمال فقدان اثر جانبی (سطح ۰) با گذشت زمان، افزایش می‌یابد و احتمالِ نداشتن اختلال معنادار با عملکرد بیمار (سطح ۱) با گذشت زمان، کاهش می‌یابد.

جدول ۱.۱۸: احتمال‌های نمونه‌ای برای سطوح مختلف $Y_۱$ و $Y_۲$ و $Y_۳$ با استفاده از پاسخ‌های مشاهده‌شده

سطح	$Y_۱$	$Y_۲$	$Y_۳$
۰	۰/۴۲۸	۰/۵۳۳	۰/٦۴۸
۱	۰/۴۲۸	۰/۳۸۲	۰/۳۲۰
۲	۰/۱۴۴	۰/۰۸۵	۰/۰۳۲

جدول ۱.۱۹ احتمال حضور در زمان t $(R_t = ۱)$ و اندازه‌ی نمونه‌ی مشاهده‌شده در سطوح مختلف پاسخ قبلی را به‌ترتیب در زمان‌های دوم و سوم $(t = ۲, ۳)$ نشان می‌دهد. جدول ۱.۱۹ نشان می‌دهد که با احتمال زیاد (۰/۹۳۵ برای $Y_۱ = ۰$ و ۰/۹۵۹ برای $Y_۱ = ۱$)، بیمارانی که دچار اثر جانبی کمتری در زمان اول شده‌اند، در زمان دوم در مطالعه باقی می‌مانند. همچنین بیمارانی که دچار اثر جانبی کمتری در زمان دوم شده‌اند، در زمان سوم با احتمال‌های زیادی (۰/۸۸۴ برای $Y_۲ = ۰$ و ۰/۸۵۹ برای $Y_۲ = ۱$) در مطالعه باقی می‌مانند. اما احتمال‌های حضور در زمان حال، هنگامی که اثر جانبی در پاسخ قبلی زیاد است، کاهش می‌یابد (۰/٦۳۴ برای $R_۲ = ۱$ زمانی که $Y_۱ = ۲$ است و ۰/۴۰۹ برای $R_۳ = ۱$ زمانی که $Y_۲ = ۲$ است).

از جدول ۱.۱۹ می‌توان نتیجه گرفت که احتمال پاسخ در زمان حال به پاسخ در زمان گذشته وابسته است. بالاتر بودن سطح پاسخ قبلی منجر به زیاد بودن احتمال انصراف می‌شود. از این احتمال‌ها می‌توان در تحلیل این داده‌ها با استفاده از معادلات براوردگر تعمیم‌یافته‌ی وزنی (فصل ۷ را ببینید) استفاده کرد.

جدول ۱۹.۱: احتمال‌های پاسخ، $R_t = 1$، اندازه‌ی نمونه‌ی مشاهده‌شده (تعداد) برای دوره‌های $t = 2$ و $t = 3$ به‌شرط سطح پاسخ‌های قبلی

Y_1	$R_2 = 1 \mid Y_1$	تعداد	Y_2	$R_3 = 1 \mid Y_2$	تعداد
۰	۰/۹۳۵	۱۱۵	۰	۰/۸۸۴	۱۲۲
۱	۰/۹۵۹	۱۱۸	۱	۰/۸۵۹	۸۵
۲	۰/۶۳۴	۲۶	۲	۰/۴۰۹	۹

۲.۵.۲.۱ کاربردی از مدل انتقالی با پاسخ ترتیبی، به‌منظور شناخت اثرهای درمانی مختلف روی داده‌های بی‌خوابی

در این مثال، مجدداً از داده‌های بی‌خوابی استفاده شده است. جدول ۲۰.۱ که از فرانکوم و دیگران (۱۹۸۹) گرفته شده است، پاسخ‌ها را به‌صورت کمتر از ۲۰ دقیقه، ۲۰—۳۰ دقیقه، ۳۰—۶۰ دقیقه و بیش‌تر از ۶۰ دقیقه طبقه‌بندی کرده است.

جدول ۲۱.۱ توزیع‌های حاشیه‌ای نمونه‌ای را برای پاسخ‌های اولیه و پیگیر برای دو نوع درمان نشان می‌دهد. از این جدول می‌توان نتایج مشابهی همچون جدول ۱۵.۱ گرفت.

جدول ۲۰.۱: مدت زمان به خواب رفتن (بر حسب دقیقه) که از پرسیدن سؤال «به چه سرعت خواب‌تان می‌برد» به دست می‌آید (پاسخ در زمان دوم، Y_2، بر اساس نوع درمان، و پاسخ در زمان اول، Y_1، تعداد مشاهده‌شده و درصد سطری)

		پاسخ در زمان دوم (Y_2)				
نوع درمان	پاسخ در زمان اول (Y_1)	< ۲۰	۲۰—۳۰	۳۰—۶۰	> ۶۰	مجموع
مؤثر	< ۲۰	۷	۴	۱	۰	۱۲
	درصد	۵۸/۳	۳۳/۳	۸/۳	۰/۰	۱۰۰/۰
	۲۰—۳۰	۱۱	۵	۲	۲	۲۰
	درصد	۵۵/۰	۲۵/۰	۱۰/۰	۱۰/۰	۱۰۰/۰
	۳۰—۶۰	۱۳	۲۳	۳	۱	۴۰
	درصد	۳۲/۵	۵۷/۵	۷/۵	۲/۵	۱۰۰/۰
	> ۶۰	۹	۱۷	۱۳	۸	۴۷
	درصد	۱۹/۱	۳۶/۲	۲۷/۷	۱۷/۰	۱۰۰/۰
دارونما	< ۲۰	۷	۴	۲	۱	۱۴
		۵۰/۰	۲۸/۶	۱۴/۳	۷/۱	۱۰۰/۰
	۲۰—۳۰	۱۴	۵	۱	۰	۲۰
	درصد	۷۰/۰	۲۵/۰	۵/۰	۰/۰	۱۰۰/۰
	۳۰—۶۰	۶	۹	۱۸	۲	۳۵
		۱۷/۱	۲۵/۷	۵۱/۴	۵/۷	۱۰۰/۰
	> ۶۰	۴	۱۱	۱۴	۲۲	۵۱
	درصد	۷/۸	۲۱/۶	۲۷/۵	۴۳/۱	۱۰۰/۰

جدول ۲۱.۱: توزیع حاشیه‌ای نمونه‌ای برای پاسخ‌های اولیه و پیگیر برای دو روش درمان

		طبقه‌بندی پاسخ			
پاسخ	نوع درمان	< ۲۰	۲۰—۳۰	۳۰—۶۰	> ۶۰
پاسخ اولیه	درمان مؤثر	۰/۱۰۱	۰/۱۶۸	۰/۳۳۶	۰/۳۹۵
	دارونما	۰/۱۱۷	۰/۱۶۷	۰/۲۹۲	۰/۴۲۵
پاسخ پیگیر	درمان مؤثر	۰/۳۳۶	۰/۴۱۲	۰/۱۶۰	۰/۰۹۲
	دارونما	۰/۲۵۸	۰/۲۴۲	۰/۲۹۲	۰/۲۰۸

۶.۲.۱ داده‌های طولی با پاسخ اسمی

در این زیربخش، یک کاربرد واقعی در مطالعات طولی با پاسخ اسمی آورده می‌شود.

۱.۶.۲.۱ بررسی تغییر وضع فعالیت اقتصادی افراد در دو فصل یکسان (پاییز) از دو سال ۱۳۸۵ و ۱۳۸۶

مثالی از مطالعات طولی اسمی می‌تواند بررسی تغییر وضع فعالیت اقتصادی افراد در طول زمان باشد. آمارگیری پانلی نیروی کار مرکز آمار ایران از جمله‌ی آمارگیری‌هایی است که می‌توان در این قسمت به کار برد. جدول ۲۲.۱ تغییر وضع فعالیت کلّیه‌ی افراد مشترک در دو فصل یکسان (پاییز) از دو سال ۱۳۸۵ و ۱۳۸۶ در طرح آمارگیری نیروی کار را نشان می‌هد. نتایج جدول ۲۲.۱ احتمال انتقال یا تغییر وضع فعالیت اقتصادی افراد مشترک بین فصل‌های پاییز سال‌های ۱۳۸۵ و ۱۳۸۶ را نشان می‌دهد. نتایج نشان می‌دهد که نسبت غیر فعالان در پاییز ۱۳۸۵ که در پاییز ۱۳۸۶ نیز غیر فعال مانده‌اند، ۸۸/۹۳ درصد است؛ در حالی که ۸/۸۸ و ۲/۱۹ درصد افراد غیر فعال در پاییز ۱۳۸۵، در پاییز ۱۳۸۶ به‌ترتیب شاغل و بیکار شده‌اند. همین تغییر وضع فعالیت در دو سال متوالی برای شاغلان در پاییز ۱۳۸۵ نیز وجود دارد، به‌گونه‌ای که ۸۲/۰۱ درصد شاغلان در پاییز ۱۳۸۵، در پاییز ۱۳۸۶ نیز شاغل مانده‌اند؛ در حالی که ۱۴/۵۱ و ۳/۴۸ درصد افراد شاغل در پاییز ۱۳۸۵، در پاییز ۱۳۸۶ به‌ترتیب غیر فعال و بیکار شده‌اند. در مورد بیکاری، با توجه به این‌که تغییر وضع فعالیت در دو سال متوالی مورد بررسی قرار گرفته است، نسبت کمی از افراد بیکار در پاییز ۱۳۸۵ (۲۶/۱۱ درصد) در سال ۱۳۸۶ بیکار باقی مانده‌اند و نسبت بیش‌تری غیر فعال (۳۱/۷۸ درصد) یا شاغل (۴۲/۱۰ درصد) شده‌اند.

با توجه به این‌که الگوی تغییر وضع فعالیت زنان و مردان متفاوت است، نتایج جدول ۲۲.۱ به تفکیک جنس نیز بیان می‌شود. نتایج جدول‌های ۲۳.۱ و ۲۴.۱ نشان‌دهنده‌ی متفاوت بودن الگوی تغییر وضع فعالیت زنان و مردان در دو دوره‌ی متوالی است. با مقایسه‌ی نتایج این دو جدول می‌توان دریافت که احتمال تحرک یا انتقال بیکاری به غیر فعالی در بین زنان (۵۸/۱۷ درصد) خیلی بیش‌تر از احتمال متناظر مردان (۲۰/۹۷ درصد) است. به‌عبارت دیگر، زنان بیکار، زودتر از مردان بیکار، از جستجوی کار، دلسرد و غیر فعال

می‌شوند. نتایج همچنین نشان می‌دهد که احتمال غیر فعال باقی ماندن زنان (۹۲/۱۴ درصد) خیلی بیش‌تر از مردان (۸۱/۵۸ درصد) است. احتمال شاغل باقی ماندن مردان (۸۶/۹۷ درصد) نیز خیلی بیش‌تر از احتمال متناظر در بین زنان (۶۱/۶۱ درصد) است.

جدول ۲۲.۱: تغییر وضع فعالیت افراد در دو فصل یکسان (پاییز) از دو سال ۱۳۸۵ و ۱۳۸۶ (تعداد، درصد سطری و ستونی)

وضع فعالیت در پاییز ۱۳۸۵	وضع فعالیت در پاییز ۱۳۸۶			
	غیر فعال	شاغل	بیکار	
غیر فعال	۲۷۴۸۸ ۸۸/۹۳ ۸۹/۵۹	۲۷۴۵ ۸/۸۸ ۱۵/۲۹	۶۷۸ ۲/۱۹ ۳۷/۱۵	۳۰۹۱۱ ۱۰۰/۰
شاغل	۲۵۳۷ ۱۴/۵۱ ۸/۲۷	۱۴۳۴۰ ۸۲/۰۱ ۷۹/۸۷	۶۰۸ ۳/۴۸ ۳۳/۳۲	۱۷۴۸۵ ۱۰۰/۰
بیکار	۶۵۶ ۳۱/۷۸ ۲/۱۴	۸۶۹ ۴۲/۱۰ ۴/۸۴	۵۳۹ ۲۶/۱۱ ۲۹/۵۳	۲۰۶۴ ۱۰۰/۰
مجموع	۳۰۶۸۱ ۱۰۰/۰	۱۷۹۵۴ ۱۰۰/۰	۱۸۲۵ ۱۰۰/۰	۵۰۴۶۰ ۱۰۰/۰

جدول ۲۳.۱: تغییر وضع فعالیت مردان در دو فصل یکسان (پاییز) از دو سال ۱۳۸۵ و ۱۳۸۶ (تعداد، درصد سطری و ستونی)

وضع فعالیت در پاییز ۱۳۸۵	وضع فعالیت در پاییز ۱۳۸۶			
	غیر فعال	شاغل	بیکار	
غیر فعال	۷۶۷۶ ۸۱/۵۸ ۸۲/۸۲	۱۴۱۱ ۱۵/۰۰ ۹/۷۸	۳۲۲ ۳/۴۲ ۲۵/۸۸	۹۴۰۹ ۱۰۰/۰
شاغل	۱۲۸۵ ۹/۱۴ ۸/۲۷	۱۲۲۳۳ ۸۶/۹۷ ۷۹/۸۷	۵۴۷ ۳/۸۹ ۳۳/۳۲	۱۴۰۶۵ ۱۰۰/۰
بیکار	۳۰۷ ۲۰/۹۷ ۳/۳۱	۷۸۲ ۵۳/۴۲ ۵/۴۲	۳۷۵ ۲۵/۶۱ ۳۰/۱۴	۱۴۶۴ ۱۰۰/۰
مجموع	۹۲۶۸ ۱۰۰/۰	۱۴۴۲۶ ۱۰۰/۰	۱۲۴۴ ۱۰۰/۰	۲۴۹۳۸ ۱۰۰/۰

بیش‌ترین احتمال تغییر وضع فعالیت مردان بیکار، به شاغل شدن آن‌ها برمی‌گردد (۵۳/۴۲ درصد) در حالی که در مورد زنان، بیش‌ترین احتمال تغییر وضع فعالیت زنان بیکار، به غیر فعال شدن برمی‌گردد (۵۸/۱۷ درصد). به‌طور کلّی می‌توان گفت که تحرک نیروی کار یا تغییر وضع فعالیت افراد برای زنان و مردان در ایران، الگوی متفاوتی دارد و این مسئله باید در تحلیل‌ها در نظر گرفته شود.

جدول ۲۴.۱: تغییر وضع فعالیت زنان در دو فصل یکسان (پاییز) از دو سال ۱۳۸۵ و ۱۳۸٦ (تعداد، درصد سطری و ستونی)

وضع فعالیت در پاییز ۱۳۸۵	وضع فعالیت در پاییز ۱۳۸٦			
	غیر فعال	شاغل	بیکار	
غیر فعال	۱۹۸۱۲ ۹۲/۱۴ ۹۲/۵۲	۱۳۳۴ ٦/۲۰ ۳۷/۸۱	۳۵٦ ۱/٦٦ ٦۱/۲۷	۲۱۵۰۲ ۱۰۰/۰
شاغل	۱۲۵۲ ۳٦/٦۱ ۵/۸۵	۲۱۰۷ ٦۱/٦۱ ۵۹/۷۲	٦۱ ۱/۷۸ ۱۰/۵۰	۳۴۲۰ ۱۰۰/۰
بیکار	۳۴۹ ۵۸/۱۷ ۱/٦۳	۸۷ ۱۴/۵۰ ۲/۴۷	۱٦۴ ۲۷/۳۳ ۲۸/۲۳	٦۰۰ ۱۰۰/۰
مجموع	۲۱۴۱۳ ۱۰۰/۰	۳۵۲۸ ۱۰۰/۰	۵۸۱ ۱۰۰/۰	۲۵۵۲۲ ۱۰۰/۰

۳.۱ دستورهای R برای محاسبه‌ی شاخص‌های مقایسه‌ای

در این بخش با استفاده از دستورهای مورد استفاده در نرم‌افزار R به محاسبه‌ی مخاطره‌ی نسبی، اختلاف نسبت‌ها، تحلیل آزمون خی‌دو و آزمون دقیق فیشر می‌پردازیم. به‌منظور بررسی اثر دود تنباکو بر ایجاد تومور مغزی در موش‌ها در مثال ۴.۱.۲.۱، با استفاده از دستورهای مربوط به آزمون خی‌دو و آزمون دقیق فیشر در نرم‌افزار R، به این سؤال پاسخ می‌دهیم. در ابتدا با استفاده از دستور matrix، داده‌های مربوط به موش‌ها در زیربخش ۴.۱.۲.۱ را می‌خوانیم و سپس با استفاده از دستورهای chisq.test و fisher.test به بررسی استقلال دو متغیر «در معرض دود تنباکو قرار گرفتن» و «وضعیت تومور» می‌پردازیم.

```
x = matrix(c(21, 19, 2, 13), ncol = 2)
chisq.test(x)
fisher.test(x, alternative = "greater")
```

در دستور fisher.test، از شناسه‌ی alternative استفاده شده است که فرض‌های مقابل $\theta > 1$ و $\theta < 1$ و $\theta \neq 1$ را به‌ترتیب با استفاده از alternative=greater و alternative=less و alternative=two.sided در نظر گیرد و مورد بررسی قرار می‌دهد. نتایج دستورهای فوق در شکل ۴.۱ نمایش داده شده است. با

```
chisq.test(x)

        Pearson's Chi-squared test with Yates' continuity correction

data:  x
X-squared = 5.3625, df = 1, p-value = 0.02057

fisher.test(x, alternative = "greater")

        Fisher's Exact Test for Count Data

data:  x
p-value = 0.008344
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 1.602124       Inf
sample estimates:
odds ratio
 6.949584
```

شکل ۴.۱: نتایج آزمون خی‌دو و آزمون دقیق فیشر برای جدول پیشایندی دود تنباکو

```
chisq.test(x)$observed
     [,1] [,2]
[1,]   21    2
[2,]   19   13

chisq.test(x)$expected
         [,1]     [,2]
[1,] 16.72727 6.272727
[2,] 23.27273 8.727273

chisq.test(x)$residuals
           [,1]      [,2]
[1,]  1.0447024 -1.705992
[2,] -0.8856896  1.446325
```

شکل ۵.۱: نتایج مربوط به مقادیر مشاهده‌شده، مورد انتظار و مانده‌های پی‌یرسون

استفاده از دستورهای chisq.test(x)$observed و chisq.test(x)$expected و chisq.test(x)$residuals می‌توان به مقادیر مشاهده‌شده، مقادیر مورد انتظار و مانده‌های پی‌یرسون [که به‌صورت $\text{(observed - expected)}/\sqrt{\text{expected}}$ تعریف می‌شود] دست یافت. این نتایج در شکل ۵.۱ داده شده است.

به‌منظور بررسی و آزمون فرض برابری نسبت‌ها، از دستور prop.test در نرم‌افزار R استفاده می‌شود. از داده‌های فلایس (۱۹۸۱، ص. ۱۳۹) برای اجرای آزمون فرض یکسان بودن نسبت سیگاری‌ها در چهار جامعه از بیماران یک بیمارستان، استفاده شده است. هدف، اجرای آزمون $H_۰ : \pi_۱ = \pi_۲ = \pi_۳ = \pi_۴$ در مقابل این فرض است که لااقل در یکی از این جامعه‌ها، این نسبت متفاوت است. در ابتدا تعداد بیماران و تعداد سیگاری‌ها در چهار جامعه را وارد می‌کنیم و سپس از دستور prop.test استفاده می‌کنیم. دستورهای

```
smokers  = c( 83, 90, 129, 70 )
```

و

```
patients = c( 86, 93, 136, 82 )
```

نشان می‌دهند که از ۸۶ فرد گروه اول، ۸۳ نفر سیگاری، از ۹۳ فرد گروه دوم ۹۰ نفر سیگاری، از ۱۳۶ فرد گروه سوم ۱۲۹ نفر سیگاری و از ۸۲ فرد گروه چهارم ۷۰ نفر سیگاری بوده‌اند. دستور

```
prop.test(smokers, patients)
```

آزمون یکسان بودن نسبت سیگاری‌ها در چهار جامعه را انجام می‌دهد. نتایج این آزمون در شکل ۶.۱ آمده است. P-مقدار در نتایج شکل ۶.۱ نشان می‌دهد که نسبت سیگاری‌ها لااقل در یکی از چهار جامعه‌ی بیماران متفاوت است و بنا بر این $H_۰$ رد می‌شود.

```
prop.test(smokers, patients)

 4-sample test for equality of proportions without continuity correction

data:  smokers out of patients
X-squared = 12.6004, df = 3, p-value = 0.005585
alternative hypothesis: two.sided
sample estimates:
   prop 1    prop 2    prop 3    prop 4
   0.965     0.967     0.948     0.853
```

شکل ۶.۱: نتایج آزمون برابری نسبت سیگاری‌ها در چهار جامعه‌ی بیماران

برای محاسبه‌ی مخاطره‌ی نسبی، باید از تقسیم دو نسبت در نرم‌افزار R استفاده کرد. به‌عنوان مثال، بر اساس اطلاعات جدول ۱۶.۱، نسبت افرادی که در کمتر از ۳۰ دقیقه به خواب می‌روند، ۰/۲۶۹ و این نسبت برای افرادی که از دارونما استفاده می‌کنند ۰/۲۸۳ است. برای محاسبه‌ی مخاطره‌ی نسبی نمونه‌ای از دستورهای زیر استفاده می‌کنیم:

```
p1 = 0.269

p2 = 0.283

rr = p1/p2.
```

۴.۱ تمرین‌ها

۱– آزمودنی‌هایی بر حسب جنس و این‌که آیا افراد کوررنگی دارند یا خیر، به‌صورت جدول ۲۵.۱ رده‌بندی شده‌اند.

جدول ۲۵.۱: وضعیت بینایی افراد به تفکیک جنسیت

	مرد	زن
طبیعی	۴۴۲	۵۱۴
کوررنگی	۳۸	۶

آیا می‌توان پذیرفت که این داده‌ها از جدول فراوانی‌های نسبی زیر (که در آن $p = 1 - q$) تبعیت می‌کنند؟ آزمون را در سطح ۰/۰۵ انجام دهید.

جدول ۲۶.۱: فراوانی‌های نسبی وضعیت بینایی افراد به تفکیک جنسیت

	مرد	زن
طبیعی	$\frac{p}{۲}$	$\frac{p^۲}{۲}+pq$
کوررنگی	$\frac{q}{۲}$	$\frac{q^۲}{۲}$

۲- جدول ۲۷.۱ را در نظر بگیرید.

جدول ۲۷.۱: مشاهدات مربوط به جدول پیشایندی ۲ × ۲ حاصل از یک متغیر پاسخ و یک متغیر کمکی دوحالتی

	$Y=۱$	$Y=۲$
$X=۱$	۲۰۰	۳۰۰
$X=۲$	۳۰۰	۲۰۰

آ) با استفاده از آماره‌ی $\chi^۲$، فرض $H_۰: p_{ij}=\frac{۱}{۴}$ را در مقابل این فرض که لااقل برای یک i و j داریم $H_۰: p_{ij}\neq\frac{۱}{۴}$ با فرض توزیع چندجمله‌ای برای تعداد در خانه‌ها آزمون کنید.

ب) آزمون بالا را با استفاده از آماره‌ی نسبت درست‌نمایی تعمیم‌یافته انجام دهید.

پ) یک فاصله‌ی اطمینان ۰/۹۵ برای نسبت بخت‌ها بیابید.

ت) فرض استقلال را با استفاده از آزمون نسبت درست‌نمایی تعمیم‌یافته آزمون کنید.

ث) اگر حاشیه‌های سطری، داده‌شده فرض شوند، چگونه استقلال را آزمون می‌کنید؟

۳- در مطالعه‌ای به بررسی تأثیر میزان مصرف سیگار با سه رده‌ی «غیر سیگاری، مصرف روزانه ۱-۱۰ عدد، مصرف روزانه بیش از ۱۰ عدد» بر نوع بیماری (۰: محدود، ۱: پیشرفته) علاقه‌مندیم. نتایج بدین صورت اعلام گردیده که نسبت بخت‌های بیماری پیشرفته برای افراد غیر سیگاری در مقابل افراد سیگاری که روزانه ۱-۱۰ عدد سیگار مصرف می‌کنند، برابر با ۰/۶۵ است و این نسبت بخت‌ها برای افراد غیرسیگاری در مقابل افرادی که روزانه بیش از ۱۰ عدد سیگار مصرف می‌کنند، برابر با ۰/۲۵ است. بر این اساس، میزان نسبت بخت‌های بیماری پیشرفته را برای افراد سیگاری با مصرف ۱-۱۰ عدد سیگار در روز، در مقابل افراد سیگاری با مصرف بیش از ۱۰ عدد سیگار در روز محاسبه کنید.

۴- فرض کنید که احتمال سالانه‌ی براورد شده‌ی مرگ بر اثر سرطان ریه

برای زنان سیگاری دارای سن بیش از ۳۵ سال ۰/۰۰۱۳۰۴ و برای غیرسیگاری‌ها ۰/۰۰۰۱۲۱ است.

آ) اختلاف نسبت‌ها و مخاطره‌ی نسبی را محاسبه و تعبیر کنید. کدام‌یک از این اندازه‌ها برای این داده‌ها بیش‌تر آگاهی‌بخش است؟ چرا؟

ب) یک فاصله‌ی اطمینان ۰/۹۵ برای مخاطره‌ی نسبی بیابید.

پ) نسبت بخت‌ها را محاسبه و تعبیر کنید. توضیح دهید که چرا مخاطره‌ی نسبی و نسبت بخت‌ها مقادیر مشابه دارند.

۵– جدول ۲۸.۱ نشان‌دهنده‌ی نتایج مطالعه‌ای آینده‌نگر بر روی موش‌های بیمار است که در آن، ۷ موش را در معرض دارویی استاندارد و ۲۰ موش را در معرض داروی جدید قرار داده‌اند. فراوانی‌های مرگ به‌صورت جدول ۲۸.۱ است.

جدول ۲۸.۱: وضعیت سلامت موش‌های تحت درمان دو نوع دارو

	زنده (۰)	مرده (۱)	جمع
داروی استاندارد (۰)	۲	۵	۷
داروی جدید (۱)	۱۶	۴	۲۰
جمع	۱۸	۹	۲۷

آ) براوردهای مخاطره‌ی نسبی، لگاریتم مخاطره‌ی نسبی، و اختلاف نسبت‌ها را برای این داده‌ها محاسبه کنید و با استفاده از روش δ، فاصله‌ی اطمینان ۹۵ درصدی برای آن‌ها بیابید.

ب) یول (۱۹۰۰) آماره‌ی زیر را برای بررسی پیوند در یک جدول پیشایندی ۲×۲ معرفی کرد:

$$Q = \frac{\pi_{۱۱}\pi_{۲۲} - \pi_{۱۲}\pi_{۲۱}}{\pi_{۱۱}\pi_{۲۲} + \pi_{۱۲}\pi_{۲۱}}.$$

نشان دهید که رابطه‌ی میان این آماره و لگاریتم نسبت بخت‌ها، γ، به‌صورت زیر است:

$$Q = \frac{e^{\gamma} - ۱}{e^{\gamma} + ۱}.$$

با استفاده از واریانس مجانبی γ (از روش دلتا)نشان دهید که واریانس مجانبی آماره‌ی Q با استفاده از روش دلتا برابر با $\frac{\sigma^{۲}}{n}$ است که در آن، $\sigma^{۲} = \frac{(\sum_i \sum_j \pi_{ij}^{-۱})(۱-Q^{۲})^{۲}}{۴}$.

(راهنمایی: از پیوست ۱ استفاده کنید.)

۶- جدول ۲۹.۱، وجود بیماری قلبی را بر اساس میزان مصرف سیگار در روز نشان می‌دهد:

جدول ۲۹.۱: وضعیت سلامت قلبی افراد بر اساس میزان مصرف سیگار در روز

	تعداد سیگار مصرف‌شده			
	۰	۱–۲۴	+۲۵	جمع
بیمار	۰	۱	۱	۲
سالم	۵	۲	۲	۹
جمع	۵	۳	۳	۱۱

استقلال میان مصرف سیگار و بیماری قلبی را با استفاده از دو روش زیر آزمون کنید و نتایج را با یکدیگر مقایسه نمایید.

آ) آزمون دقیق (توزیع فوق‌هندسی)،

ب) آزمون مجانبی χ^2 پی‌یرسون.

۷- بر اساس اطلاعات جدول ۱۴.۱ نشان دهید که آیا نوع درمان بر پاسخ زمان اول و زمان دوم به‌طور جداگانه بستگی دارد یا نه؟ آماره‌ی d سامرز را برای بررسی همبستگی بین متغیرهای نوع درمان و مدت زمان به خواب رفتن در زمان اول، و نوع درمان و مدت زمان به خواب رفتن در زمان دوم، با استفاده از نرم‌افزار SAS یا SPSS به دست آورید.

۸- بر اساس اطلاعات جدول ۲۰.۱، با استفاده از چه معیاری می‌توان پیوند بین پاسخ‌های ترتیبی زمان‌های اول و دوم را به دست آورد؟ با استفاده از نرم‌افزارهای آماری، برای درمان مؤثر و استفاده از دارونما به‌طور مجزا، آماره‌ی مناسب برای بررسی پیوند بین پاسخ‌ها را بیابید. همچنین با استفاده از اطلاعات جدول ۲۱.۱، نشان دهید که آیا بین پاسخ زمان اول و نوع درمان، پیوندی وجود دارد یا نه. آماره‌ای مناسب برای بررسی پیوند بین پاسخ زمان دوم و متغیر کمکی نوع درمان معرفی کنید و با استفاده از یک نرم‌افزار مناسب، آن را محاسبه کنید.

۹- برای جدول ۳۰.۱ که در آن، متغیر Y دارای سه سطح a و b و c است و متغیر X دارای دو سطح d و e است، توزیع دقیق $\chi^۲$ را تحت فرض استقلال بیابید. اگر جدول ۳۱.۱ مشاهده شده باشد، P-مقدار را برای فرض استقلال X و Y بیابید (به اگرستی، ۲۰۰۲، صص.۵۴-۵۶ رجوع کنید).

جدول ۳۰.۱: اطلاعات حاشیه‌ای مربوط به جدول پیشایندی حاصل از یک متغیر پاسخ و یک متغیر کمکی رده‌بندی‌شده

			Y		
		a	b	c	
X	d				۲
	e				۵
		۱	۲	۴	۷

جدول ۳۱.۱: مشاهدات مربوط به جدول پیشایندی حاصل از یک متغیر پاسخ و یک متغیر کمکی رده‌بندی‌شده

			Y		
		a	b	c	
X	d	۱	۱	۰	۲
	e	۰	۱	۴	۵
		۱	۲	۴	۷

۱۰- برای $i = ۱, ۲, \ldots, N$ گیریم Y_i به‌شرط x_i دارای توزیع برنولی با احتمال موفقیت زیر است:

$$p_i(\theta) = \Pr(Y_i = ۱|x_i) = \frac{\exp(\alpha + \theta x_i)}{۱ + \exp(\alpha + \theta x_i)}.$$

آ) نشان دهید که اگر $H_۰ : \theta = ۰$ و $H_۱ : \theta = \beta$ باشد، برای مقادیر معلوم α و β، آزمون نسبت درست‌نمایی، $H_۰ : \theta = ۰$ را به‌ازای مقادیر بزرگ آماره‌ی $\beta \sum x_i Y_i$ رد می‌کند.

ب) آماره‌ی آزمون فوق را استاندارد کنید و نشان دهید که برای Nهای بزرگ، آماره‌ی مناسب برای آزمون $H_۰ : \theta = ۰$ در مقابل $H_۱ : \theta \neq ۰$ عبارت است از

$$T(Y) = \frac{\sum_i x_i(Y_i - p)}{\{p(۱ - p)\sum_i x_i^۲\}^۲},$$

که در آن $p = \frac{e^\alpha}{۱+e^\alpha}$.

پ) برای مقادیر θ نزدیک به صفر، نشان دهید که

$$p_i(\theta) = p + \theta x_i(۱ - p).$$

ت) تحت دنباله‌ی آزمون‌های $\theta = \theta_N = \frac{\lambda}{N^{\frac{1}{2}}}$ (تحت H_1) میانگین $T(Y)$ را یافته، نشان دهید که واریانس آن برای Nهای بزرگ، نزدیک به ۱ است.

۱۱– برای جدول زیر، همه‌ی مقادیر ممکن χ^2 را برای آزمون $H_0 : \theta = 1$ در مقابل $H_1 : \theta > 1$ محاسبه کرده، توزیع دقیق χ^2 را به دست آورید. همچنین همه‌ی مقادیر ممکن P–مقدار را محاسبه کنید و میانگین P–مقدار را به دست آورید.

جدول ۳۲.۱: اطلاعات حاشیه‌ای مربوط به جدول پیشایندی ۲ × ۲ حاصل از یک متغیر پاسخ و یک متغیر کمکی دوحالتی

۲		
۲		
	۲	۲

۱۲– فرض کنید برای $j = 1, 2, \ldots$ داشته باشیم $\Pr(T = t_j) = \pi_j$. نشان دهید که $E[\text{مقدار}–(\text{mid}–P)] = 0/5$.

(راهنمایی: نشان دهید که $\frac{(\sum_j \pi_j)^2}{2} = (\sum_j \pi_j(\frac{\pi_j}{2} + \pi_{j+1} + \cdots)$)

۱۳– در جدول زیر، تعداد اتومبیل‌های بیمه‌شده توسط شرکت‌های بیمه‌ی مختلف ارائه شده است که در طی یک ماه پس از انعقاد قرارداد، دچار سانحه گشته‌اند. مدل پواسونی به این داده‌ها برازش دهید و نیکویی برازش آن را با استفاده از آماره‌ی خی‌دو آزمون کنید.

جدول ۳۳.۱: تعداد اتومبیل‌های بیمه‌شده‌ی دچار سانحه

تعداد دفعات اتفاق حادثه	فراوانی مشاهده‌شده
۰	۱۴۴
۱	۳۰
۲	۲۰
۳+	۵

۱۴– توزیع چندجمله‌ای با K رده را در نظر بگیرید. فرض کنید p_i احتمال قرار گرفتن در رده‌ی iام باشند و n_i تعداد مشاهدات که در این رده قرار می‌گیرند (از n مشاهده‌ی ممکن). برای آزمون $H_0 : p_i = p_{i0}$ برای مقادیر مشخص‌شده‌ی $p_{10}, \ldots, p_{K0}$ در مقابل این فرض که لااقل برای یک i داریم $H_1 : p_i \neq p_{i0}$، نشان دهید که

$$-2\ln\lambda = 2\sum_{i=1}^{K} O_i \ln\left(\frac{O_i}{E_i}\right),$$

که در آن $O_i = n_i$، $E_i = np_{i\circ}$، $(i = ۱, ۲, \ldots, K)$، و λ آماره‌ی آزمون ماکسیمم درست‌نمایی تعمیم‌یافته‌ی $H_\circ$ در مقابل $H_۱$ است.

۱۵- فرض کنید پژوهشگری با استفاده از احتمال خطای نوع اول برابر با ۰/۰۵، فرض صفر را رد می‌کند اگر P-مقدار از ۰/۰۵ کمتر یا مساوی با آن باشد. فرض کنید یک آزمون دقیق که از $\chi^۲$ استفاده می‌کند، تحت فرض صفر دارای توزیع زیر باشد:

۹	۳	۰	a
۰/۰۸	۰/۶۲	۰/۳	$\Pr(\chi^۲ = a)$

آ) نشان دهید که با استفاده از P-مقدار، احتمال واقعی خطای نوع اوّل برابر با صفر است.
ب) نشان دهید برای mid-P، مقدار احتمال واقعی خطای نوع اوّل، برابر با ۰/۰۸ است.
ت) با استفاده از احتمال‌های ۰/۳ و ۰/۶۶ و ۰/۰۴ به ترتیب برای مقادیر ۰ و ۳ و ۹ به جای احتمال‌های خانه‌های جدول بالا، احتمال واقعی خطای نوع اوّل را با P-مقدار و (mid-P)-مقدار به دست آورید. P-مقدار محافظه‌کارانه‌تر است یا (mid-P)-مقدار؟ چرا؟