



دانشکده بهداشت

دفتر توسعه آموزش

طرح دوره (Course Plan)

نام درس: تحلیل داده های طبقه بندی شده	کد درس:
مقطع تحصیلی: کارشناسی ارشد	رشته تحصیلی: آمارزیستی
گروه آموزشی: آمارزیستی	پیش نیاز:
تعداد واحد: ۳ واحد	محل برگزاری: دانشکده بهداشت
روز و ساعت برگزاری: شنبه ۱۲-۱۰ و دو شنبه ۱۴-۱۲	مسئول درس: دکتر آزاده ساکی
☐ طراحی اولیه ☒ بازنگری	
هدف کلی دوره: تحلیل داده های رسته ای در علوم پزشکی مبتنی بر روش و هدف مطالعه	
اهداف اختصاصی دوره: ۱- مفاهیم و تعاریف داده های گسسته و انواع آن را تعریف کنند. ۲- توزیع احتمال مربوط به انواع داده های گسسته شامل دوجمله ای، چندجمله ای و پواسن و دوجمله ای منفی در انواع مطالعات را تشخیص دهند. ۳- ارتباط سنجی و آزمون های مرتبط با پاسخ های رسته ای را بدون استفاده از مدل سازی مطابق با هدف مطالعه انجام دهند. ۴- مدل های خطی تعمیم یافته را تعریف کنند و توابع پیوند مناسب با هر نوع داده را استخراج و مدلسازی با کد نویس در نرم افزار انجام دهند. ۵- مدل های متداول برای پاسخ های دودویی شامل لجستیک، پروبیت، خطی را تعریف و بتوانند با استفاده از نرم افزار مدلسازی را انجام دهند. ۶- مدل رگرسیون لجستیک را بر روی داده های مثال برازش داده و ضرائب آن را تفسیر و توضیح دهند. ۷- آزمون نیکویی برازش مدل و انتخاب بهترین مدل در لجستیک چندگانه را انجام دهند. ۸- مدل لوگ خطی را بر روی داده های شمارشی بر ارزش داده و تفسیر کنند و مفاهیم آن را توضیح دهند. ۹- آزمون نیکویی برازش مدل لگ خطی و انتخاب بهترین مدل را انجام دهند. ۱۰- بیش پراکنی در مدل های گسسته را شناسایی و علل آن را توضیح دهند. ۱۱- بدون استفاده از مدل نسبت شانس و بررسی ارتباط را در مطالعه مورد - شاهدهی همسان سازی شده محاسبه کنند. ۱۲- مدل های حاشیه ای ، GEE برای پاسخ های دودویی همبسته را تعریف و مدلسازی را با استفاده از نرم افزار انجام دهند.	
جدول ترتیب ارائه محتوای آموزشی در طول دوره	
جلسه	رئوس مطالب
اول	تعریف داده های گسسته و توزیع های احتمال آن ها - استنباط آماری برای سهم ها و توابع درستمایی و آزمون های مرتبط با آن - معرفی آزمون های دقیق ، آزمون جایگشت، آزمون فیشر، و تکنیک Bootstrap



دانشکده بهداشت

دفتر توسعه آموزش

طرح دوره (Course Plan)

دوم	جداول توافقی و ساختار احتمال آن ها - توزیعهای احتمال متناسب با جداول توافقی در انواع مطالعات مشاهده ای- کاربرد جداول توافقی در بررسی ارتباط دو متغیر کیفی-محاسبه نسبت شانس و محاسبه حساسیت و اختصاصیت- آزمون مقایسه نسبت در دو جامعه مستقل-
سوم	بررسی مخدوشگرها در جداول توافقی ، نسبت شانس شری و حاشیه ای، همگنی نسبت شانس در جداول توافقی- آزمون متل هنزل و پارادوکس سیمپسون
چهارم	استفاده از نرم افزار برای تحلیل سهم ها و جداول توافقی - حل تمرینهای انتخاب شده توسط دانشجویان
پنجم	معرفی مدل های خطی تعمیم یافته مفاهیم ومولفه های مدل های خطی تعمیم یافته - تعریف تابع پیوند و انواع آن- معرفی مدل های مناسب برای داده های دو حالتی ، چند حالتی، رتبه ای ، و شمارشی
ششم	چک کردن مناسبت مدل ها بر اساس باقیمانده های مدل، توابع پیشگویی کننده و مناسبت تابع پیوند
هفتم	معرفی تابع glm در R و انجام تمرینهای منتخب
هشتم	معرفی مدل رگرسیون لجستیک و رگرسیون لجستیک چندگانه و برازش آن روی داده های حقیقی و تفسیر پارامترهای مدل
نهم	استنباط آماری برای مدل رگرسیون لجستیک و بررسی مناسب مدل و انتخاب بهترین مدل بر اساس معیارهای AIC و AICC و بررسی بیش پراکنش روش های ساخت و به کارگیری یک مدل رگرسیون لجستیک مناسب روش پیشرو، پس رو و گام به گام -آزمون های مناسب بر اساس شاخصهای LR , $deviance$
دهم	بررسی قدرت پیشگویی مدل بر اساس جداول طبقه بندی، منحنی راک و همبستگی روش های تعیین حجم نمونه مناسب برای برازش مدل لجستیک
یازدهم	تحلیل مدل رگرسیون لجستیک چندگانه و بررسی مناسب مدل با R و انجام تمرین های منتخب
دوازدهم	مدل های لجستیک برای داده های کیفی چند جمله ای و معرفی مدل های لجستیک تجمعی و لجستیک رتبه ای برای داده های کیفی رتبه ای و برازش و تفسیر نتایج مدل
سیزدهم	تحلیل مدل های رگرسیون لجستیک برای پاسخ های رتبه ای و بررسی مناسب مدل ها با R و انجام تمرین های منتخب
چهاردهم	حل تمرین و رفع اشکال
پانزدهم	معرفی مدل لوگ خطی برای داده های شمارشی و جداول توافقی دو راهه و سه راهه
شانزدهم	استنباط در مورد پارامترهای مدل و بررسی ارتباطات شرطی و غیر شرطی، ادغام جداول و استفاده از رگرسیون لجستیک، بررسی مناسب مدل و بیش پراکنشی در این مدل ها
هفدهم	تحلیل مدل لوگ خطی برای پاسخ های شمارشی و بررسی مناسب مدل ها با R و انجام تمرین های منتخب
هجدهم	حل تمرین و رفع اشکال
نوزدهم	تحلیل مطالعات همسان سازی شده، و آزمون مک نمار برای بررسی استقلال در مطالعات همسان سازی شده
بیستم	محاسبه odds ratio در مطالعات همسان سازی شده



دانشکده بهداشت

دفتر توسعه آموزش

طرح دوره (Course Plan)

بیستم و یکم	استفاده از نرم افزار R برای تحلیل جداول توافقی در مطالعات همسان شده و برآورد OR - حل تمرینهای انتخاب شده توسط دانشجویان
بیست و دوم	مدل رگرسیون لجستیک برای مطالعات همسان سازی شده - معرفی مدل های حاشیه ای و مدل های شرطی
بیست و سوم	استفاده از نرم افزار R و برازش مدل های حاشیه ای و شرطی در مطالعات همسان شده و پاسخهای همبسته - حل تمرینهای انتخاب شده توسط دانشجویان
بیست و چهارم	حل تمرین و رفع اشکال
بیست و پنجم	مرور مطالب و جمع بندی
روش های آموزشی (نحوه ارائه درس): PBL، بحث گروهی، سخنرانی و اسلاید، ارایه سناریو - تفکر انتقادی در حل تمرینها	
وسایل آموزشی: کامپیوتر - دیتا - کتاب	
وظایف و تکالیف دانشجوی: مطالعه مباحث و جستجو در پایگاههای معتبر برای هر جلسه. حل تمرینهای منتخب و ارایه در کلاس	
روش های ارزیابی دانشجوی (لطفا سهم هر روش بر حسب درصد نوشته شود): روش های ارزیابی در طول دوره: پاسخ به پرسشهای مطرح شده در کلاس ۲۰٪ - حل تمرینهای منتخب ۲۰٪ - حضور فعال ۶۰٪ - روش های ارزیابی در انتهای دوره: آزمون پایان دوره ۶۰٪	
منابع اصلی درس:	
۱- Agresti, A. An introduction to categorical data analysis. Third Ed.2019	
۲- Collett, D Modeling Binary Data. Second Ed.2002	
۳- Hosmer, D. W. and Lemeshow, S. Applied Logistic Regression. Second ED.2000	
۴- Fleiss, J. I. statistical methods for rates and proportions, second Ed.1981	