

تحلیل داده‌های طبقه‌بندی شده

تعداد واحد: ۳

نوع واحد: نظری

پیشنیاز: روشهای آمار زیستی (۱) و (۲)

کد درس: ۱۵

هدف: بررسی روشهای تجزیه و تحلیل آماری وقتی که صفت وابسته بصورت کیفی باشد.

سرفصل درس: (۵۱ ساعت)

استنباط آماری برای داده‌های دودوئی؛ مدل‌های آماری برای داده‌های دودوئی و دو جمله‌ای؛ چک کردن مدل؛ بیش-پراکنش؛ مدل‌های آمیخته برای داده‌های دودوئی

رئوس مطالب:

- استنباط آماری برای داده‌های دودوئی شامل: توزیع دوجمله‌ای؛ استنباط راجع به احتمال موفقیت؛ مقایسه دو نسبت؛ مقایسه چند نسبت.
- روشهای تجزیه و تحلیل داده‌های دو بعدی و سه بعدی بدون مدل‌بندی
- مدل‌های آماری برای داده‌های دودوئی و دو جمله‌ای شامل: کلیات مدل سازی آماری؛ مقدمه‌ای بر مدل‌های خطی؛ روشهای برآورد کردن؛ مدل لژستیک خطی؛ روند خطی در نسبت‌ها
- چک کردن مدل‌ها شامل: تعریف باقیمانده‌ها؛ چک کردن فرم پیش‌گوئی‌کننده‌های خطی؛ چک کردن کفایت تابع ربط
- بیش پراکنش شامل: علل بالقوه بیش پراکنش؛ مدل‌سازی تغییرات در احتمالات پاسخ؛ مدل‌سازی همبستگی بین پاسخ‌های دودوئی
- مدل‌های آمیخته برای داده‌های دودوئی شامل: تعریف اثر ثابت و اثر تصادفی؛ مدل‌های آمیخته؛ مدل‌سازی چند سطحی؛ مدل‌های آمیخته برای داده‌های طولی
- روشهای دقیق شامل: مقایسه دو نسبت با استفاده از آزمون دقیق؛ رگرسیون لژستیک دقیق برای یک پارامتر؛ آزمون فرضیه دقیق
- مدل‌های خطی تعمیم‌یافته، رگرسیون لژستیک، رگرسیون پواسن، مدل‌های لوجیت

منابع اصلی درس:

- 1- Collett, D. *Modeling Binary Data*. second Ed. 2002
- 2- Hosmer, D. W. and Lemeshow, S. *Applied Logistic Regression*. second Ed. 2000
- 3- Agresti, A. *Categorical Data Analysis*. second Ed. 2002

شیوه ارزیابی دانشجویان:

در این درس دانشجویان به صورت تراکمی (امتحان پایان ترم) و تکوینی (کار در خلال ترم، میان ترم، پروژه و غیره) ارزیابی خواهند شد. درصد امتحان پایان ترم ۶۰٪، کار ترمی از قبیل انجام تکالیف و حضور کلاسی ۲۰٪ و امتحان میان ترم ۲۰٪ پیشنهاد می‌شود.

