



آزمون ناپارامتری Kolmogorov-Smirnov (KS)

مقدمه:

آزمون ناپارامتری KS برای آزمودن یک توزیع آماری دلخواه در مقایسه با توزیع آماری در آزمایش تجربی دیگر و یا یک توزیع مرجع بکار میرود. داده های بزرگ با فرض استقلال داده ها و یگانگی توزیع فراوانی آنها تابع CLT خواهند بود اما در بررسی داده کوچک و حتی داده بزرگ با عدم وجود شرط استقلال داده ها، به آزمون ناپارامتری نیاز خواهیم داشت. نیز اغلب در فعالیتهای تحقیقی و در مراحل اولیه آنها (هنگامیکه اطلاعات بسیار محدود است) آزمون ناپارامتری لازم خواهد بود.

هدف اولیه پروژه، آشنایی با توابع متلب/R، برنامه نویسی و تحقیق یک تست ناپارامتری بر روی داده واقعی است. موارد مورد نیاز از مرجع درس و مراجع معتبر دیگر مطالعه شود.

شرح پروژه:

۱. آزمون KS را در سطح اعتبار ۵٪ با فرض Null توزیع یکنواخت $[0,1]$ بر روی داده ذیل (X) انجام دهید:
 $X=[0.58, 0.42, 0.52, 0.33, 0.43, 0.23, 0.58, 0.76, 0.53, 0.64]$

۲. دو مجموعه داده توسط یک توزیع دلخواه (با پارامترهای متفاوت) ایجاد کنید و آزمون KS را در سطح اعتبار ۱٪ بر روی آنها انجام دهید.

۳. آزمون KS را در سطح اعتبار ۵٪ و ۱٪ با فرض Null توزیع یکسان بر روی داده های پیوست انجام دهید. p-value در دو آزمون چیست؟

۴. نتایج آزمون KS را با نتایج آزمون Student's T مقایسه کنید. دلایل شکست آزمون Student's T چه مواردی میتواند باشد؟

ارایه شما شامل موارد بالا بعلاوه یک مستند گزارش پروژه (حداقل ۵۰۰ کلمه و حداکثر ۱۰۰۰ کلمه) میبایست از طریق درس افزار (CourseWare) بارگذاری شود. تمام فایل های لازم برای اجرا به همراه کدهای خود را بصورت فشرده (zip) و بطور جداگانه بارگذاری کنید. ضمن گزارش عملکرد و نتایج پروژه، گزارش کنید که آزمون های دیگر در این زمینه چیست؟ (حداقل ۲ مورد) و چه تفاوت هایی میان آزمون KS و این آزمون ها وجود دارد؟ چرا آزمون KS موثرتر از آزمون Student's T است؟

انجام پروژه:

برای امکان بررسی یکسان، برنامه به متلب یا R خواهد بود. استفاده از سایر زبان های برنامه نویسی و یا نظایر آن، مشمول ۵۰٪ (حداقل) جریمه خواهد شد. هرگونه خطای برنامه نویسی و یا اجرا بر روی کامپیوتر منتخب (توسط بررسی کننده کد) مشمول جریمه خواهد بود. بنابراین شما باید مطمئن شوید که تمام فایل های مورد نیاز ارسال شده و نیز کد شما عاری از هرگونه خطای احتمالی باشد. همچنین کامنت گذاری بسیار اهمیت خواهد داشت.

GUI برای مجموعه عملیات و مقایسه خروجیها ضروری نیست اما ۲۵٪ امتیاز اضافه خواهد داشت. همچنین رتبه اول بهترین پروژه، ۲۵٪ و رتبه دوم ۱۲.۵٪ امتیاز اضافه دریافت خواهند کرد.

داده ها:

داده ها در دو فایل اکسل جداگانه به پیوست این راهنما در درس افزار موجود است.

دقت کنید که طول داده ها یکسان نیست اما بهر حال این مساله مشکلی در آزمون KS ایجاد نمیکند!